

EINE METHODE ZUR PRÜFUNG
MEHRDIMENSIONALER HYPOTHESEN
IM NICHTPARAMETRISCHEN FALL

Christina Schneider

EINE METHODE ZUR PRÜFUNG
MEHRDIMENSIONALER HYPOTHESEN
IM NICHTPARAMETRISCHEN FALL

Inaugural-Dissertation
zur Erlangung des Grades Doctor rerum politicarum
(Dr.rer.pol.)
an der Ludwig-Maximilians-Universität München

vorgelegt von
Christina Schneider

1984

Referent: Prof. Dr. Kurt Weichselberger

Korreferent: Prof. Dr. Franz Ferschl

Promotions-

abschlußberatung: 20. Februar 1985



Digitized by the Internet Archive
in 2026

G l i e d e r u n g

	Seite
§ 1 Einleitung	1
1.1 Bisherige Entwicklungen auf dem Gebiet der Statistically Equivalent Blocks	1
1.2 Ein Beispiel zur Konstruktion von Statistically Equivalent Blocks	9
1.3 Konstruktion von Statistically Equivalent Blocks	16
1.4 Gegenstand der Arbeit	29
§ 2 Konsistenz und Statistically Equivalent Blocks	31
2.1 Partitionenfolgen, die ein Wahrscheinlichkeitsmaß eindeutig festlegen	34
2.2 Folgen von Partitionen und Statistically Equivalent Blocks	45
§ 3 Definition der Statistically Equivalent Blocks durch Abbildungen	53
3.1 Vorbereitende Definitionen	53
3.2 Definition einer Unterabbildung in Anlehnung an den eindimensionalen Fall	
3.3 Differenzierbarkeit der Abbildungen F und F^{-1}	63
3.4 Betrachtungen bei festem Stichprobenumfang	71
3.5 Definition einer Metrik	80

	Seite
§ 4 Vorschlag einer Testgröße	92
4.1 Bemerkungen zur Dirichlet-Ver- teilung	95
4.2 Die Testgröße	97
4.3 Die Verteilung der Testgröße	98
4.4 Konsistenz des Tests	107
§ 5 Abschließende Bemerkungen	109
Anhang zu § 4	112
Literaturverzeichnis	114

Die vorliegende Arbeit entstand während meiner Tätigkeit als wissenschaftliche Angestellte am Institut für Statistik und Wissenschaftstheorie der Ludwig-Maximilians-Universität München.

Herrn Professor Dr. K. Weichselberger möchte ich für sein die Arbeit begleitendes Interesse, seine Unterstützung und seine kritischen Hinweise danken.

Im besonderen gilt mein Dank auch Herrn Professor Dr. F. Ferschl für die anregenden und hilfreichen Gespräche.

Bei Frau Miko Hohmann und Frau Sylvia Hintz möchte ich mich für die geduldige und sorgfältige Anfertigung des Maschinenmanuskripts bedanken.

Häufig verwendete Symbole

Die folgende Aufzählung beinhaltet die Bezeichnungen, die in der vorliegenden Arbeit häufig verwendet werden. Die Seitenzahl hinter diesen Symbolen gibt an, wo sich zum erstenmal eine Definition des genannten Begriffes findet.

	Seite		Seite
$B_{O,k_1}(\vec{X})$	17	$A(E), A_O(E)$	35
U_{k_1}	17	$H = (h_1, \dots, h_n)$	42
Z_{k_1}	17	$h_i, i = 1, \dots, n$	42
$\underline{k}_i, \bar{k}_i, i = 1, \dots, n$	21	$(\underbrace{\mathbb{R}_p \times \dots \times \mathbb{R}_p}_{n\text{-mal}})_O$	56
$B_{\underline{k}_i, \bar{k}_i}(\vec{X}), i = 1, \dots, n$	21	$S_\pi((\vec{X}_1, \dots, \vec{X}_n))$	57
$Z_{k_i}, i = 2, \dots, n$	21	$F((z_{k_1}, \dots, z_{k_n}))$	57
$U_{k_i}, i = 2, \dots, n$	23	$F_i((z_{k_1}, \dots, z_{k_i})),$ $i = 1, \dots, n$	57
$V_{k_i}, i = 1, \dots, n$	23	$F^{-1}((u_{x_1}, \dots, u_{x_n}))$	59
$\vec{V} = (V_1, \dots, V_n)$	25	$F_i^{-1}((u_{x_1}, \dots, u_{x_i})),$ $i = 1, \dots, n$	59
$B_i(\vec{X}), i = 1, \dots, n+1$	26	Tr_μ	60
$(P : n \in \mathbb{N})$	35		
(μ) -Partitionenfolgen	35		

	Seite		Seite
$\underbrace{(\text{Tr}_u \times \dots \times \text{Tr}_u)}_{n\text{-mal}} \circ$	60	F_p	81
D_π	60	$d_\circ$	82
D_+	78	d_2	86
F_+	78	$F^n(\sigma(x))$	100

Häufig verwendet wird auch das Symbol $B_{\underline{k}_i, \bar{k}_i}(\vec{x})$,
 $i = 1, \dots, n$ und $B_i(\vec{x})$, $i = 1, \dots, n+1$. Das " $\vec{x}$ " hinter
dem Mengensymbol deutet lediglich an, daß es sich
hier um *nichtzufällige* Größen handelt!

Wie üblich bezeichnet B_p die σ -Algebra der Borelschen
Mengen auf $\mathbb{R}_p$, $p \geq 2$ und λ_p das Lebesgue - Borelsche
Maß auf $(\mathbb{R}_p; B_p)$.

Unter $\mathbb{R}_+$ ist die Menge der positiven reellen Zahlen
zu verstehen. Sei A eine Menge, so bezeichnet $|A|$ die
Mächtigkeit dieser Menge und $\bar{A}$ deren Komplement bzgl.
einer fest vorgegebenen Obermenge.

Sei $A \subset E$, so bezeichnet $I_A: E \longrightarrow \{0,1\}$ die Indika-
torfunktion dieser Menge

$$I_A(x) = \begin{cases} 1 & \text{falls } x \in A \\ 0 & \text{falls } x \in \bar{A} \end{cases}$$

Für die identische Abbildung einer Menge auf sich selbst wird das Symbol "id" verwendet.

Unter pr_i ist die Projektion der i -ten Koordinaten von $\mathbb{R}_p$, $i \leq p$ auf $\mathbb{R}$ zu verstehen:

$$\begin{aligned} pr_i: \mathbb{R}_p &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_p) &\longmapsto x_i, \quad 1 \leq i \leq p \end{aligned}$$

Mit Γ wird die Gammafunktion bezeichnet.

Auf Seite 100 befindet sich das Symbol

$F^n(g(x))$. Es bezeichnet - Existenz vorausgesetzt - eine der Funktionen G , für die gilt:

$$\frac{d^n}{dx^n} G(x) = g(x).$$

Unter einer *Dirichlet* - Verteilung ist, wenn nichts anderes bemerkt wird, der Einfachheit halber immer eine *Dirichlet* - Verteilung zu verstehen, deren Parameter alle identisch 1 sind.

§ 1

Einleitung

1.1 Bisherige Entwicklungen auf dem Gebiet der Statistically Equivalent Blocks

Unter *Statistically Equivalent Blocks* versteht man eine endliche und zufällige Zerlegung des Raumes $\mathbb{R}_p$, $p \geq 2$. Diese Partitionen sind aus einer Stichprobe n unabhängiger, stetiger und identisch verteilter Zufallsvektoren auf $(\mathbb{R}_p, \mathcal{B}_p, \mu)$ konstruiert worden wobei μ die zugrunde liegende Verteilung bezeichne. Für jede Realisation der Stichprobe - bei zugrunde gelegtem Stichprobenumfang - erhält man eine "feste" Zerlegung in $n + 1$ in $\mathcal{B}_p$ enthaltenen Teilmengen. Integriert man die zufälligen Mengen mit dem zugrundeliegenden Wahrscheinlichkeitsmaß μ und faßt n dieser $n + 1$ Zufallsvariablen zu einem Zufallsvektor $(V_1, \dots, V_n)$ zusammen, so ist dieser, unabhängig von der speziellen Gestalt des vorgegebenen stetigen Wahrscheinlichkeitsmaßes μ , *Dirichlet* - verteilt mit den Parametern $\theta_0 = \theta_1 = \dots = \theta_n = 1$. Für die einzelnen Komponenten gilt¹⁾:

$$E(V_i) = \frac{1}{n+1}, \quad i = 1, \dots, n;$$

hierin dürfte auch die Begründung für die Bezeichnung "equivalent" liegen.

¹⁾ siehe hierzu näheres S. 95

Mit dem Zufallsvektor $(V_1, \dots, V_n)$ hat man eine Übertragung der Stichprobe, deren mehrdimensionale Verteilung oft unhandlich ist, in eine Zufallsgröße mit zugänglicher Verteilung erreicht. Oft ist es günstiger an Stelle von $(V_1, \dots, V_n)$ den Zufallsvektor

$$(U_1, \dots, U_n),$$

$$U_i = \sum_{j=1}^i V_j, \quad i = 1, \dots, n$$

zu betrachten.

Dieser ist dann verteilt wie die Orderstatistik n unabhängiger, auf dem Intervall $[0;1]$ gleichverteilter Zufallsvariabler. Die Unabhängigkeit von der speziellen Gestalt der zugrundeliegenden Verteilung - nur die Steitigkeit ist vorauszusetzen - bedingt die Anwendbarkeit dieser Methode in mehrdimensionalen, nichtparametrischen Fragestellungen. Insbesondere ist es in Testsituationen möglich, auf die "transformierten" Zufallsvariablen bekannte Tests, wie z.B. den *Kolmogorov*-*Smirnov* - oder *Cramér* - *v.Mises* - Test anzuwenden.

Auch wenn man die Entstehung dieser Konstruktionsweise betrachtet, stand immer eine nichtparametrische Fragestellung im Hintergrund; so war ein erster Schritt in der Entwicklung dieses Verfahrens die Konstruktion eines p -dimensionalen, $p \geq 2$, nichtparametrischen Toleranzbereichs. Im eindimensionalen Fall ist dieser ein Intervall und die zu seiner Konstruktion zu lösende Aufgabe lautet:

Ausgehend von einer Stichprobe unabhängiger und identisch verteilter Zufallsvariablen $X_1, \dots, X_n$ und zu vorgegebenem $\alpha \in (0;1)$ und $\beta \in (0;1)$ finde man zwei reelle Funktionen der Stichprobe

$$U((X_1, \dots, X_n)) \text{ und } O((X_1, \dots, X_n)) ,$$

so daß

$$W(W_X([U((X_1, \dots, X_n)); O((X_1, \dots, X_n))]) : \alpha) = \beta$$

gilt.

Wilks¹⁾ löste 1941 dieses Problem für *stetige* Verteilungsannahmen. Die obere und untere Grenze dieses Intervalls waren hierbei geeignete Orderstatistiken $X_{(r)}, X_{(s)}$, $0 < r < s \leq n$ der Stichprobe und die Verteilung der Zufallsvariablen

$$W_X([X_{(r)}; X_{(s)}])$$

- unabhängig von der speziellen Gestalt der zugrundeliegenden Verteilung - eine *Beta*-Verteilung, deren Parameter nur vom Rang der ermittelten Orderstatistiken abhängen.

A. Wald²⁾ verallgemeinerte diese Vorgehensweise auf p -dimensionale, stetige Verteilungsannahmen.

Ausgehend von einer Stichprobe n , $n \in \mathbb{N}$ unabhängiger stetiger und identisch verteilter Zufallsvektoren auf $(\mathbb{R}_p, \mathcal{B}_p, \mu)$ konstruiert er einen zufälligen Quader Q

1) Wilks, S. S.: Determination of Sample Sizes for Setting Tolerance Limits. Ann. math. Statistics 12, 91 - 96 (1941).

2) Wald A.: An Extension of Wilks' Method for Setting Tolerance Limits. Ann. math. Statistics 14, 45 - 55 (1943).

- als direkte Verallgemeinerung von Wilks' Vorgehen -, dessen Eckpunkte aus Koordinaten bestehen, die Orderstatistiken der Stichproben der einzelnen Koordinaten der p -dimensionalen Stichprobe sind. Zur Herleitung der Verteilung von $\mu(Q)$ sei $n \geq 2p$ vorausgesetzt; er beginnt die r_1 -te und s_1 -te Orderstatistik Z_{r_1} und Z_{s_1} der Stichprobe der ersten Komponenten zu ermitteln, so daß $s_1 > r_1$ und $s_1 - r_1 - 1 \geq 2(p - 1)$ gilt;

hiermit erhält er die Menge

$\text{pr}_1^{-1}((Z_{r_1}; Z_{s_1}])$; von den $s_1 - r_1 - 1$ hierin liegenden Stichprobenelementen betrachtet er die zweiten Komponenten und ermittelt hiervon die s_2 -te und r_2 -te Orderstatistik Z_{r_2}, Z_{s_2} , wobei

$s_2 > r_2$ und $s_2 - r_2 - 1 \geq 2(p - 2)$ gelte.

Dies ergibt die Menge

$$\text{pr}_1^{-1}((Z_{r_1}, Z_{s_1}]) \cap \text{pr}_2^{-1}((Z_{r_2}, Z_{s_2}]) ;$$

so fährt er, in jedem Schritt die "nächsten" Komponenten verarbeitend fort, bis er nach p Schritten, entsprechende Nebenbedingungen als erfüllt vorausgesetzt, den Quader

$$Q = \bigcap_{i=1}^p \text{pr}_i^{-1}((Z_{r_i}; Z_{s_i}])$$

erhalten hat. Diese Menge ist zufällig und A. Wald¹⁾ konnte die Verteilung von $\mu(Q)$ bestimmen, eine *Beta*-Verteilung, deren Parameter nur von s_p und r_p abhängen.

¹⁾ siehe Seite 3 unter 2)

Die Verteilung dieser Zufallsgröße ist also abhängig von der Reihenfolge, in der die einzelnen Komponenten verarbeitet werden. Mit dieser Menge wurden $2p$ weitere, disjunkte, zufällige Mengen konstruiert, denen Wald in dieser Arbeit jedoch keine Beachtung schenkte. Dies verarbeitete aber J. W. Tukey, der Abraham Walds Vorgehen in zweifacher Hinsicht ausbaute und verallgemeinerte und somit die Möglichkeit wies, das Verfahren als "Transformation" komplizierter p -dimensionaler Verteilungsannahmen in Dirichlet-Verteilungen entsprechender Parameter zu verwenden. Die Verallgemeinerungen waren im einzelnen:

- es werden nicht nur $2p + 1$ sondern $n + 1$ disjunkte Teilmengen erzeugt
- zugelassen werden nicht nur Projektionen auf die Koordinatenachsen, sondern n , nicht notwendig verschiedene reelle Abbildungen $h_1, \dots, h_n$, die nur zu gewährleisten haben, daß die $h_i(\vec{X})$ $i = 1, \dots, n$ wieder stetige Zufallsvariablen sind.

Er beginnt damit, ausgehend von einer Stichprobe n unabhängiger, stetiger und identisch verteilter Zufallsvektoren, diese mit der Funktion h_1 zu transformieren und das größte der Bilder Z_n zu ermitteln: hiermit erhält er die Mengen

$$h_1^{-1}((-\infty; Z_n]) \text{ und } h_1^{-1}((Z_n; \infty))$$

die verbleibenden $n - 1$ Stichprobenelemente, die in $h_1^{-1}((-\infty; Z_n])$ liegen, transformiert er mittels h_2 und bestimmt hieraus wiederum das größte Element Z_{n-1} , und erhält die Menge

1) Tukey, J. W.: Non Parametric Estimation II, Statistically Equivalent Blocks and Tolerance Regions. The Continuous Case. Ann. math. Statistics 18, 529 - 539 (1947).

$$h_2^{-1}((-\infty; z_{n-1}]) \cap h_1^{-1}((-\infty; z_n]) \text{ und}$$

$$h_2^{-1}((z_{n-1}, \infty)) \cap h_1^{-1}((-\infty, z_n]).$$

Dieses Vorgehen wird n-mal wiederholt, bis alle Stichprobenelemente "aufgebraucht" sind und führt schließlich zu $n + 1$ disjunkten und zufälligen Mengen

$$B_1, \dots, B_{n+1}.$$

Die Bildung der einzelnen Mengen unterliegt bei ihm noch einer fest vorgegebenen Reihenfolge.

Sei μ das zugrundeliegende Wahrscheinlichkeitsmaß, so kann aus jeweils n dieser Mengen der Zufallsvektor

$$\vec{V} := (V_1, \dots, V_n) := (\mu(B_1), \dots, \mu(B_n))$$

gebildet werden.

J. W. Tukey war es möglich, die gemeinsame Verteilung, die Randverteilungen sowie die bedingten Verteilungen von $(V_1, \dots, V_n)$ zu bestimmen; von ihm stammt auch die seitdem für die Mengen B_i gebräuchliche Bezeichnung *Statistically Equivalent Blocks*.¹⁾ 1951 wies D. A. S. Fraser²⁾ in einer Arbeit darauf

¹⁾ "Statistically Equivalent Blocks" wird im weiteren Verlauf der Arbeit immer mit der Bezeichnung "StEB's" abgekürzt werden.

²⁾ Fraser, D. A. S.: Sequentially determined Statistically Equivalent Blocks. Ann. math. Statistics 20, 372 - 381 (1951).

hin, daß es nicht notwendig ist, um die Verteilungsgesetze zu erhalten, die vorgegebenen Schnittfunktionen $h_1, \dots, h_n$ sukzessive zu "verarbeiten"; vielmehr ist es möglich, die Wahl der nächsten Schnittfunktion h_k von den vorangegangenen $h_1, \dots, h_{k-1}$ abhängig zu machen oder sie durch ein Zufallsexperiment aus einer Menge vorgegebener, für das Verfahren geeigneter Funktionen auszuwählen. Erstere Verfahrensweise kann vor allem dann Bedeutung erlangen, wenn man z.B. bei der Bestimmung von Toleranzbereichen eine Vorstellung von der geometrischen Form dieser Menge hat.

Zwei Jahre später, 1953, formuliert er das Verfahren neu¹⁾, in der allgemeinsten Form, die auch T. W. Anderson²⁾ in seiner Darstellung übernommen hat. Da Andersons Darstellung wohl die eleganteste ist, wird hier nicht auf Frasers Originalfassung eingegangen, sondern im nächsten Abschnitt das Verfahren in Anlehnung an T. W. Anderson referiert. Schließlich ist noch der 1956 erschienene Artikel von J. H. B. Kemperman³⁾ zu erwähnen, in dem zwar

1) Fraser, D. A. S.: Nonparametric Tolerance Regions. Ann. math. Statistics 24, 44 - 55 (1953).

2) Anderson, T. W.: Some Nonparametric Multivariate Procedures Based on Statistically Equivalent Blocks. Multivariate Analysis. Proceedings of an International Symposium held in Dayton, Ohio, 1965. Edited by P. R. Krishnaiah, 5 - 27. Academic Press, New York and London (1966).

3) Kemperman, J. H. B.: Generalized Tolerance Limits. Ann. math. Statistics 27, 180 - 186 (1956).

dem Verfahren nichts mehr hinzugefügt wird. Lediglich die Wahl der Schnittfunktionen wird in einer Weise verallgemeinert wie sie jedoch für die meisten Zwecke ohne Bedeutung sein dürfte. Die Herleitungen und Beschreibungen aller an dem Verfahren beteiligten Verteilungen findet jedoch in dieser Arbeit die durchsichtigste und klarste Darstellung. A. Wald, D. A. S. Fraser, J. W. Tukey und J. H. B. Kemperman haben in ihren Arbeiten das Verfahren auch für diskrete Verteilungsannahmen modifiziert. Da in der vorliegenden Arbeit davon nicht Gebrauch gemacht werden wird, sei auf eine Darstellung verzichtet.

1.2 Ein Beispiel zur Konstruktion von Statistically Equivalent Blocks

Das in diesem Abschnitt vorzustellende Verfahren ist eine Verallgemeinerung der Wahrscheinlichkeitstransformation stetiger, eindimensionaler Zufallsvariabler; dort erhält man, indem man eine stetige Zufallsvariable X mit ihrer Verteilungsfunktion F transformiert, die auf dem Einheitsintervall gleichverteilte Zufallsgröße $U := F(X)$. U ist das Wahrscheinlichkeitsmaß bezüglich der Verteilung μ von X eines nach links unbeschränkten Intervalls mit zufälliger oberer Grenze X :

$$U = \mu((-\infty; X]).$$

Aus einer Stichprobe n unabhängiger, identisch verteilter, stetiger Zufallsgrößen $X_1, \dots, X_n$ wird nun durch die Wahrscheinlichkeitstransformation mit ihrer Verteilungsfunktion F eine Stichprobe n unabhängiger auf dem Einheitsintervall gleichverteilter Zufallsgrößen. Die Dichte deren Orderstatistik $U_1, \dots, U_n$ lautet:

$$f((u_1, \dots, u_n)) = \begin{cases} n! & \text{falls } 0 \leq u_1 \leq \dots \leq u_n \leq 1 \\ 0 & \text{sonst.} \end{cases}$$

Geht man zu einer Stichprobe von n unabhängigen, identisch verteilten, stetigen, p -dimensionalen, $p \geq 2$, Zufallsvektoren $\vec{X}$ über und transformiert deren k -te, $1 \leq k \leq p$, Komponente X_k mit der entsprechenden Randverteilungsfunktion, so erhält man wieder n unabhängige auf dem Einheitsintervall gleichverteilte Zufallsvariable. Diese sind dann die Wahrscheinlichkeitsmaße bezüglich der zugrundeliegenden Verteilung von n linearen Halbräumen, die durch die Ungleichungen

$$\text{pr}_k(\vec{x}) \leq \text{pr}_k(\vec{x}_i) , \quad i = 1, \dots, n$$

bestimmt sind; jede andere p - dimensionale Zufallsvariable mit der gleichen k - ten Randverteilung führt nach dieser Transformation zu einer auf dem Einheitsintervall gleichverteilten Zufallsgröße. Unterstellt man nun als Zielvorstellung eine Stichprobe n unabhängiger und identisch verteilter, p - dimensionaler, $p \geq 2$, Zufallsgrößen $\vec{x}_1, \dots, \vec{x}_n$ zu n eindimensionalen Zufallsgrößen $U_{k_1}, \dots, U_{k_n}$ zu "transformieren", die zum einen eine "einfache" Verteilung besitzen mögen, wenn μ das zugrundeliegende Wahrscheinlichkeitsmaß war und zum anderen möglichst eine andere Verteilung besitzen, wenn μ nicht das zugrundeliegende Wahrscheinlichkeitsmaß war, so erfüllt die nachfolgende, als Beispiel zu verstehende Vorgehensweise die erste Anforderung; daß sie auch unter nicht allzu einschränkenden Regularitätsbedingungen die zweite Forderung erfüllt, wird in § 3 gezeigt werden. Darüberhinaus stellt sie einen Spezialfall des im Anschluß vorgestellten Verfahrens StEB's zu bilden dar.

Zu diesem Beispiel seien:

- $\vec{x}_1, \dots, \vec{x}_n$ eine Stichprobe n unabhängiger, stetiger und identisch verteilter, zweidimensionaler Zufallsvektoren
 $\vec{x}_i := (X_{1,i}, X_{2,i}) , \quad i = 1, \dots, n$
- f deren Dichte und F die Verteilungsfunktion;
 f_1, f_2 und F_1, F_2 bezeichnen die entsprechenden Randdichten und Randverteilungsfunktionen.

o.E. transformiere man die ersten Komponenten $X_{1,1}, \dots, X_{1,n}$ mit der Randverteilungsfunktion F_1 zu $F_1(X_{1,1}), \dots, F_1(X_{1,n})$ und bestimme hiervon die n - te Orderstatistik $U_n := F_1(X_{(1,n)})$.

$X_{(1,n)}$ bezeichne die hierfür verantwortliche Zufallsgröße.

U_n ist Beta-verteilt und besitzt die Dichte

$$g_1(u_n) = \begin{cases} n \cdot u_n^{n-1} & \text{falls } 0 \leq u_n \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Die verbleibenden Zufallsvektoren werden

$\vec{x}_1^{(2)}, \dots, \vec{x}_{n-1}^{(2)}$ genannt. Sie sind unter der Be-

dingung $x_{1,i}^{(2)} \leq x_{(1,n)}$ unabhängig und identisch

verteilt; ihre Verteilung besitzt die Dichte:

$$f_2^B((x_{1,i}^{(2)}, x_{2,i}^{(2)})) = \begin{cases} \frac{f((x_{1,1}^{(2)}, x_{2,1}^{(2)}))}{u_n} & \text{falls } x_{1,1}^{(2)} \leq x_{(1,n)} \\ 0 & \text{sonst} \end{cases}$$

für $i = 1, \dots, n-1$.

Transformiert man jetzt die zweiten Komponenten

$x_{2,1}^{(2)}, \dots, x_{2,n-1}^{(2)}$ mittels der Randverteilungsfunktion dieser bedingten Verteilung zu

$$y_i^{(2)} := \frac{F((x_{1,n}, x_{2,1}^{(2)}))}{u_n} \quad i=1, \dots, n-1$$

so gilt für $i=1, \dots, n-1$

$$W(Y_i^{(2)} \leq u) = \begin{cases} u & \text{falls } 0 \leq u \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Hieraus folgt, daß die Zufallsvariablen $Y_i^{(2)} \cdot u_n$ für $i=1, \dots, n-1$ unter der Bedingung $U_n = u_n$ auf dem Intervall $[0; u_n]$ gleichverteilt sind.

Wählt man weiter die $(n-1)$ -te Orderstatistik von

$Y_1^{(2)} \cdot u_n, \dots, Y_{n-1}^{(2)} \cdot u_n$ und bezeichnet diese mit U_{n-1} ,

so hat die bedingte Verteilung von U_{n-1} unter der Bedingung $U_n = u_n$ die Dichte

$$g_2^B(u_{n-1}|u_n) = \begin{cases} (n-1) \cdot \frac{u_n^{n-2}}{u_n^{n-1}} & \text{falls } 0 \leq u_{n-1} \leq u_n \\ 0 & \text{sonst} \end{cases}.$$

Die gemeinsame Verteilung von (U_{n-1}, U_n) besitzt dann die Dichte

$$g_2((u_{n-1}, u_n)) = \begin{cases} (n-1) \cdot n \cdot u_n^{n-2} & \text{falls } 0 \leq u_{n-1} \leq u_n \leq 1 \\ 0 & \text{sonst} \end{cases}.$$

Für ein "festes" (u_{n-1}, u_n) gilt dann:

$$u_{n-1} = F((x_{(1,n)}, x_{(2,n-1)}^{(2)})) \quad \text{ sowie }$$

$$u_n = F_1(x_{(1,n)}).$$

Analog der vorangehenden Überlegungen sind die verbleibenden $n-2$ Zufallsvektoren

$$\vec{x}_1^{(3)}, \dots, \vec{x}_{n-2}^{(3)} \quad \text{unter den Bedingungen}$$

$$x_{1,i}^{(3)} \leq x_{(1,n)} \quad \text{ und } \quad x_{2,i}^{(3)} \geq x_{(2,n-1)}^{(2)} \quad i=1, \dots, n-2$$

unabhängig und identisch verteilt mit der Dichte

$$f_3^B((x_{1,i}^{(3)}, x_{2,i}^{(3)})) = \begin{cases} \frac{f((x_{1,i}^{(3)}, x_{2,i}^{(3)}))}{u_{n-1}} & \text{falls } x_{1,i}^{(3)} \leq x_{(1,n)} \text{ und } x_{2,i}^{(3)} \geq x_{(2,n-1)}^{(2)} \\ 0 & \text{sonst} \end{cases}.$$

Erstellt man von dieser Stichprobe wieder, wie im ersten Schritt, die Wahrscheinlichkeitstransformation der ersten Komponente, so erhält man die Zufallsvariablen

$$Y_i^{(3)} := \frac{F((X_{1,i}^{(3)}, X_{(2,n-1)}^{(2)}))}{u_{n-1}} \quad i=1, \dots, n-2.$$

$Y_i^{(3)}$ ist unter den Bedingungen $U_{n-1} = u_{n-1}$ und $U_{n-1} \leq U_n \leq u_n$ auf dem Intervall $[0; u_{n-1}]$ gleichverteilt für $i=1, \dots, n-2$. Die $(n-2)$ -te Orderstatistik von $Y_1^{(3)} \cdot u_{n-1}, \dots, Y_{n-2}^{(3)} \cdot u_{n-1}$ sei mit $U_{n-2} := F((X_{(1,n-2)}^{(3)}, X_{(2,n-1)}^{(2)}))$ bezeichnet; sie besitzt die bedingte Dichte

$$g_3^B(u_{n-2} | u_{n-1}, u_n) = \begin{cases} (n-2) \cdot \frac{u_{n-2}^{n-2}}{u_{n-1}^{n-2}} & \text{falls } 0 \leq u_{n-2} \leq u_{n-1} \leq u_n \leq 1 \\ 0 & \text{sonst} \end{cases};$$

Für die gemeinsame Dichte ergibt sich dann

$$g_3((u_{n-2}, u_{n-1}, u_n)) = \begin{cases} (n-2)(n-1)n \cdot u_{n-2}^{n-3} & \text{falls} \\ & 0 \leq u_{n-2} \leq u_{n-1} \leq u_n \leq 1 \\ 0 & \text{sonst} \end{cases}$$

Führt man so, von den jeweils verbleibenden Zufallsvektoren alternierend die ersten und zweiten Komponenten betrachtend fort, so erhält man schließlich nach dem letzten Schritt den Zufallsvektor $(U_1, \dots, U_n)$, dessen Dichte die Gestalt

$$g_n((u_1, \dots, u_n)) = \begin{cases} n! & \text{falls } 0 \leq u_1 \leq \dots \leq u_n \leq 1 \\ 0 & \text{sonst} \end{cases}$$

besitzt. Die Zufallsvariablen lassen sich folgendermaßen schreiben:

$$U_n = F_1(X_{(1,n)}^{(1)}) = F(X_{(1,n)}^{(1)}, \infty)$$

$$U_{n-i} = \begin{cases} F((X_{(1,n-i)}^{(i+1)}, X_{(2,n-(i-1))}^{(i)})) & \text{für gerades } i \\ F((X_{(1,n-(i+1))}^{(i)}, X_{(2,n-i)}^{(i+1)})) & \text{für ungerades } i \end{cases}$$

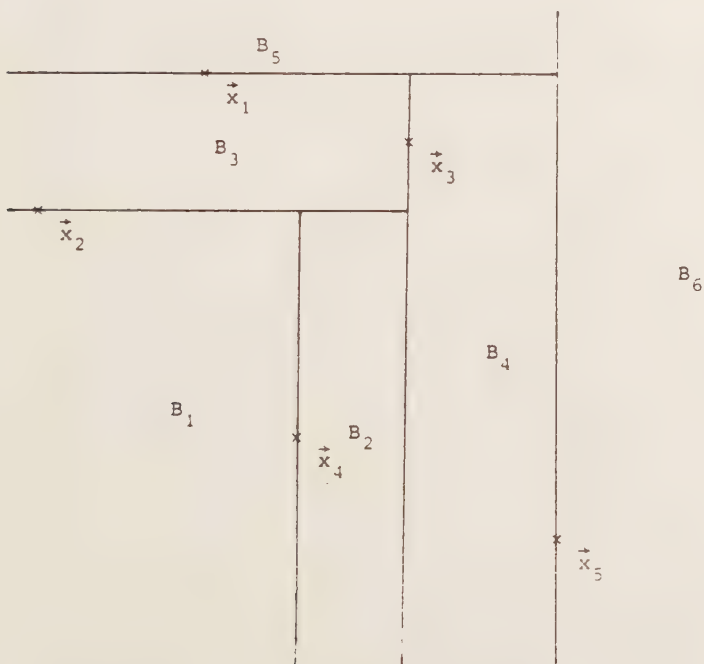
Das Verfahren StEB's zu bilden, wie T. W. Anderson¹⁾ es vorstellt, verallgemeinert die obige Vorgehensweise in folgenden Punkten:

- Für die Dimension der betrachteten Zufallsgrößen gilt: $p \geq 2$.
- Es werden nicht nur die Projektionen der Zufallsvektoren zur Bildung der StEB's herangezogen, sondern allgemein reelle Funktionen $h_1, \dots, h_n$ zugelassen, die zu gewährleisten haben, daß die Bilder $h_i(\vec{X}_j)$ der Zufallsvektoren $\vec{X}_1, \dots, \vec{X}_n$ wieder stetige und reelle Zufallsvariablen sind.
- Es wird darauf verzichtet, in jedem Schritt nur die größte der transformierten Variablen verarbeiten zu können, vielmehr wird in jedem Schritt eine "geeignete" andere Orderstatistik aus einer Teilmenge der verbleibenden Variablen ermittelt; dies bedingt die Notwendigkeit, Fallunterscheidungen zu treffen.

Daß auch unter diesen Verallgemeinerungen sinnvolle Zufallsgrößen $U_1, \dots, U_n$ entstehen, wird zum einen durch die Vorgabe einer Permutation der Indices $\{1, \dots, n\}$, deren Handhabung ein wesentlicher Bestandteil des Verfahrens ist, und zum anderen durch die Festlegung, welche der Abbildungen $h_1, \dots, h_n$ in welchem Schritt des Verfahrens Verwendung findet, gewährleistet.

¹⁾ siehe Seite 7 unter 2).

Folgende Skizze möge das Vorgehen, wie es eingangs dieses Abschnittes vorgeführt wurde, veranschaulichen. Dabei seien die $\vec{x}_1, \dots, \vec{x}_5$ Ergebnisse unabhängiger und identisch verteilter, zweidimensionaler Zufallsvektoren, die wie nachstehend eingezeichnet realisiert wurden.



Im nächsten Abschnitt wird das Verfahren, StEB's zu bilden, in der auf der vorangegangenen Seite angesprochenen, allgemeinen Form dargestellt werden.

1.3 Konstruktion von Statistically Equivalent Blocks in der Darstellung von T. W. Anderson

Bei der Darstellung StEB's zu bilden, wird die Vorgehensweise T. W. Andersons¹⁾ im wesentlichen übernommen.

Hierzu seien:

- $\vec{X}_1, \dots, \vec{X}_n$ n unabhängige, identisch verteilte Zufallsvektoren auf $(\mathbb{R}_p; \mathcal{B}_p)$, deren Verteilung mit μ bezeichnet werde und die stetige Verteilungsfunktion F_μ besitze,
- $h_1, \dots, h_n$ n nicht notwendig verschiedene, reelle Funktionen

$h_i: \mathbb{R}_p \longrightarrow \mathbb{R} \quad i=1, \dots, n$, die bewirken, daß die Zufallsvariablen

$h_i(\vec{X}_j) \quad i=1, \dots, n \text{ und } j=1, \dots, n$

wieder stetige Verteilungsfunktionen besitzen.

Im Beispiel des vorangegangenen Abschnitts waren für geradzahliges n $n/2$ dieser Funktionen die Projektion auf die erste Koordinate und $n/2$ dieser Funktionen die Projektion auf die zweite Koordinate.

- $\pi: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$
 $i \longrightarrow \pi(i)$

sei eine Permutation der Indices und es werde

$k_i := \pi(i)$ gesetzt.

Mit diesen Elementen können jetzt in n Schritten die $(n+1)$ StEB's definiert werden und mit ihnen ein n -dimensionaler Zufallsvektor

$(U_{k_1}, \dots, U_{k_n}), \quad 0 \leq U_{k_i} \leq 1, \quad i=1, \dots, n$

Im ersten Schritt transformiert man die ganze

Stichprobe $\vec{X}_1, \dots, \vec{X}_n$ mittels h_{k_1} zu $h_{k_1}(\vec{X}_1), \dots, h_{k_1}(\vec{X}_n)$

¹⁾ Siehe Seite 7 unter²⁾.

und ermittelt hiervon die k_1 -te Orderstatistik;
diese werde in diesem Fall ausnahmsweise mit
 Z_{k_1} bezeichnet.

Mit Hilfe dieser Zufallsgröße bildet man die ersten
beiden - zufälligen - Mengen

$$B_{0,k_1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) \leq Z_{k_1}\} \text{ und}$$

$$B_{k_1,n+1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) > Z_{k_1}\}$$

sowie die Zufallsvariable

$$U_{k_1} := \mu(B_{0,k_1}(\vec{X})).$$

Um die Verteilung dieser Zufallsvariable zu ermitteln,
stelle man die folgenden Überlegungen an:

$Z_i := h_{k_1}(\vec{X}_i)$, $i=1, \dots, n$ sind n unabhängige und iden-
tisch verteilte, stetige Zufallsvariable und für alle
 $i=1, \dots, n$ ist

$$U_i := \mu(\{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) \leq Z_i\}) = h_{k_1}(\mu)((-\infty; Z_i]);$$

U_i sind die Wahrscheinlichkeitstransformationen
der Z_i , $i=1, \dots, n$ und U_{k_1} bezeichne die k_1 -te
Orderstatistik dieser Zufallsvariablen; sie ist
Beta-verteilt mit den Parametern k_1 und $n-k_1+1$
und besitzt somit die Dichte

$$f_{k_1}(u_{k_1}) = \frac{n!}{(k_1-1)!(n-k_1)!} \cdot u_{k_1}^{k_1-1} (1-u_{k_1})^{n-k_1} I_{[0;1]}(u_{k_1}).$$

Da diese Zufallsvariablen stetig sind, ist die
 k_1 -te Orderstatistik fast sicher eindeutig bestimmt;
deswegen geht man davon aus, im zweiten Schritt die
verbleibenden $(n-1)$ Zufallsvektoren zu betrachten.

In die Menge $B_{0,k_1}(\vec{X})$ fallen (k_1-1) dieser Zufalls-
größen, sie seien mit $\vec{X}_1^{(2)}, \dots, \vec{X}_{k_1-1}^{(2)}$ bezeichnet.

$(n-k_1)$ Zufallsvektoren fallen in die Menge $B_{k_1,n+1}(\vec{X})$,
die mit $\vec{X}_{k_1+1}^{(2)}, \dots, \vec{X}_n^{(2)}$ bezeichnet werden.

An dieser Stelle ist für die weiteren Überlegungen
die folgende Fallunterscheidung zu treffen:

- i) $k_2 < k_1$
- ii) $k_2 > k_1$

Im ersten Fall transformiert man die k_1-1 in $B_{0,k_1}(\vec{X})$ gelegenen Zufallsvektoren mittels h_{k_2} zu $h_{k_2}(\vec{X}_1^{(2)}), \dots, h_{k_2}(\vec{X}_{k_1-1}^{(2)})$ und bezeichne hieraus die k_2 -te Orderstatistik mit Z_{k_2} . Z_{k_2} ist eine reelle Zufallsvariable. Hiermit bildet man die beiden Mengen $B_{0,k_2}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_2}(\vec{x}) \leq Z_{k_2}\} \cap B_{0,k_1}(\vec{X})$ und $B_{k_2,k_1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_2}(\vec{x}) > Z_{k_2}\} \cap B_{0,k_1}(\vec{X})$ sowie die reelle Zufallsvariable $U_{k_2} := \mu(B_{0,k_2}(\vec{X}))$.

Im zweiten Fall verfährt man analog zum ersten und transformiert die $n-k_1$ Stichprobenelemente aus $B_{k_1,n+1}(\vec{X})$ mittels der Funktion h_{k_2} . Aus den transformierten, reellen Zufallsvariablen ermittle man jetzt die k_2-k_1 -te Orderstatistik und bezeichne diese mit Z_{k_2} .

Daraus bildet man nun die Mengen

$B_{k_1,k_2}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_2}(\vec{x}) \leq Z_{k_2}\} \cap B_{k_1,n+1}(\vec{X})$ und $B_{k_2,n+1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_2}(\vec{x}) > Z_{k_2}\} \cap B_{k_1,n+1}(\vec{X})$ sowie die reelle Zufallsvariable $U_{k_2} := \mu(B_{0,k_1}(\vec{X})) + \mu(B_{k_1,k_2}(\vec{X})) = U_{k_1} + \mu(B_{k_1,k_2}(\vec{X}))$.

Im ersten Fall gilt $U_{k_1} \geq U_{k_2}$ und im zweiten Fall $U_{k_2} \geq U_{k_1}$, wobei wegen der Stetigkeit der Zufallsgrößen in beiden Fällen das Gleichheitszeichen mit Wahrscheinlichkeit 0 eintritt.

Zur Herleitung der gemeinsamen Verteilung von (U_{k_1}, U_{k_2}) betrachte man zunächst die Verteilung von U_{k_2} unter der Bedingung, daß $Z_{k_1} = z_{k_1}$ bereits realisiert wurde:

Sei $k_2 < k_1$,

$$B_{0, k_1}(\vec{x}) = \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) \leq z_{k_1}\} \quad \text{und}$$

$$B_{k_1, n+1}(\vec{x}) = \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) > z_{k_1}\} \quad \text{sowie}$$

$$u_{k_1} = \mu(B_{0, k_1}(\vec{x})), \quad \text{dann ist}$$

$$\mu_{k_2}(B) := \frac{\mu(B \cap B_{0, k_1}(\vec{x}))}{\mu(B_{0, k_1}(\vec{x}))} \quad \text{für alle } B \in \mathcal{B}_p$$

die bedingte Wahrscheinlichkeit von B unter $B_{0, k_1}(\vec{x})$.

Unter der Bedingung, daß $\vec{x} \in B_{0, k_1}(\vec{x})$ gilt, besitzt

$h_{k_2}(\vec{x})$ die stetige Verteilung $h_{k_2}(\mu_{k_2})$.

$$\text{Sei } U := \mu(B_{0, k_1}(\vec{x}) \cap \{\vec{x} \in \mathbb{R}_p : h_{k_2}(\vec{x}) \leq h_{k_2}(\vec{x})\}),$$

$$\text{so gilt } U = h_{k_2}(\mu_{k_2})((-\infty; h_{k_2}(\vec{x})) \cdot u_{k_1}.$$

Betrachtet man $h_{k_2}(\mu)((-\infty; h_{k_2}(\vec{x}))$ als Wahrscheinlichkeitstransformation von $h_{k_2}(\vec{x})$, so ergibt sich wegen der Stetigkeit von $h_{k_2}(\mu)$:

$$W(U \leq u) = W(h_{k_2}(\mu)((-\infty; h_{k_2}(\vec{x})) \leq \frac{u}{u_{k_1}}) = \frac{u}{u_{k_1}} \cdot I_{[0; u_{k_1}]}(u).$$

Die k_2 -te Orderstatistik einer Stichprobe $(k_1 - 1)$ unabhängiger auf $[0; u_{k_1}]$ gleichverteilter Zufallsvariabler besitzt die folgende Dichte:

$$f_{k_2}(u_{k_2}) = \frac{(k_1 - 1)!}{(k_2 - 1)!(k_1 - k_2 - 1)!} \cdot \frac{u_{k_2}^{k_2 - 1} (u_{k_1} - u_{k_2})^{k_1 - k_2 - 1}}{u_{k_1}^{k_1 - 1}} \cdot I_{[0; u_{k_1}]}(u_{k_2}).$$

Im Fall $k_1 < k_2$, d.h. $u_{k_1} < u_{k_2}$, ergibt sich in analoger Weise

$$f_{k_2}(u_{k_2}) = \frac{(n-k_1)!}{(k_2-k_1-1)!(n-k_2)!} \frac{(u_{k_2}-u_{k_1})^{k_2-k_1-1} (1-u_{k_2})^{n-k_2}}{(1-u_{k_1})^{n-k_1}} \cdot I_{(u_{k_1}; 1]}(u_{k_2})$$

Die gemeinsamen Dichten von $0 \leq U_{k_2} \leq U_{k_1} \leq 1$ bzw.

$0 \leq U_{k_1} \leq U_{k_2} \leq 1$ lauten dann:

$$f_{k_1, k_2}((u_{k_1}, u_{k_2})) = \frac{n!}{(k_2-1)!(k_1-k_2-1)!(n-k_1)!} \cdot u_{k_2}^{k_2-1} \cdot (u_{k_1}-u_{k_2})^{k_1-k_2-1} (1-u_{k_1})^{n-k_1}$$

für $0 \leq u_{k_2} \leq u_{k_1} \leq 1$

und

$$f_{k_1, k_2}((u_{k_1}, u_{k_2})) = \frac{n!}{(k_1-1)!(k_2-k_1-1)!(n-k_2)!} \cdot u_{k_1}^{k_1-1} \cdot (u_{k_2}-u_{k_1})^{k_2-k_1-1} (1-u_{k_2})^{n-k_2}$$

für $0 \leq u_{k_1} \leq u_{k_2} \leq 1$.

Jetzt nach dem zweiten Schritt ist der Stichprobenraum in drei disjunkte Teilmengen aufgeteilt:

$B_{0, k_2}(\vec{X})$, $B_{k_2, k_1}(\vec{X})$, $B_{k_1, n+1}(\vec{X})$ falls $k_2 < k_1$ gilt und

$B_{0, k_1}(\vec{X})$, $B_{k_1, k_2}(\vec{X})$, $B_{k_2, n+1}(\vec{X})$ falls $k_1 < k_2$ gilt.

In den einzelnen Mengen liegen jeweils k_2-1 , k_1-k_2-1 und $n-k_1$ bzw. k_1-1 , k_2-k_1-1 und $n-k_2$ der verbleibenden $n-2$ Elemente der Stichprobe.

Für die allgemeine Darstellung des i -ten Schrittes, $1 \leq i \leq n$, sei die folgende, auch weiterhin verwendete Definition getroffen:

$$\underline{k}_1 := 0, \quad \bar{k}_1 := n+1$$

$$\underline{k}_i := \max \{j \in \{0, k_1, \dots, k_{i-1}\} : j < k_i\}$$

$$\bar{k}_i := \min \{j \in \{k_1, \dots, k_{i-1}, n+1\} : j > k_i\} \quad i=2, \dots, n$$

Der Index i des Buchstaben k bedeutet: man verwende im i -ten Schritt die Abbildung h_{k_i} . Will man nun diesen Schritt durchführen, bezeichnet $\underline{k}_i$ den *größten* der schon verwendeten Indices k_j , der kleiner als k_i ist und $\bar{k}_i$ den *kleinsten* der schon verwendeten Indices, der größer als k_i ist;

sei beispielsweise $0 < k_2 < k_1 < n$, dann ist

$$\underline{k}_2 = \max \{j \in \{0, k_1\} : j < k_2\} = 0 \quad \text{und}$$

$$\bar{k}_2 = \min \{j \in \{k_1, n+1\} : j > k_2\} = k_1.$$

Im i -ten Schritt ermittelt man zunächst $\underline{k}_i$ und $\bar{k}_i$ und hiermit eine Menge $B_{\underline{k}_i, \bar{k}_i}(\vec{X})$, die durch einen der vorangegangenen Schritte entstanden ist. Diese Menge enthält, wie anschließend gezeigt werden wird, $\bar{k}_i - \underline{k}_i - 1$ der bis jetzt noch nicht "verbrauchten" Ele-

mente der Stichprobe $\vec{X}_{\underline{k}_i+1}^{(i)}, \dots, \vec{X}_{\bar{k}_i-1}^{(i)}$; diese trans-

formiert man mittels der Abbildung h_{k_i} zu

$h_{k_i}(\vec{X}_{\underline{k}_i+1}^{(i)}), \dots, h_{k_i}(\vec{X}_{\bar{k}_i-1}^{(i)})$, ermittelt hiervon die

$(k_i - \underline{k}_i)$ -te Orderstatistik und bezeichnet diese mit Z_{k_i} . Hiermit werden die beiden folgenden Mengen definiert:

$$B_{\underline{k}_i, k_i}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq Z_{k_i}\} \cap B_{\underline{k}_i, \bar{k}_i}(\vec{X}) \quad \text{und}$$

$$B_{k_i, \bar{k}_i}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) > Z_{k_i}\} \cap B_{\underline{k}_i, \bar{k}_i}(\vec{X}).$$

Die Menge $B_{\underline{k}_i, k_i}(\vec{X})$ enthält $k_i - \underline{k}_i - 1$ und die Menge

$B_{k_i, \bar{k}_i}(\vec{X})$ $\bar{k}_i - k_i - 1$ der verbleibenden Stichprobenelemente.

Die obigen Definitionen zusammenfassend, können jetzt

die durch das Verfahren entstehenden Mengen definiert werden:

$$B_{0,n+1}(\vec{X}) := \mathbb{R}_p$$

$$B_{0,k_1}(\vec{X}) := \{\vec{X} \in \mathbb{R}_p : h_{k_1}(\vec{X}) \leq z_{k_1}\}$$

$$B_{k_1,n+1}(\vec{X}) := \{\vec{X} \in \mathbb{R}_p : h_{k_1}(\vec{X}) > z_{k_1}\}$$

$$B_{\underline{k}_i, k_i}(\vec{X}) := \{\vec{X} \in \mathbb{R}_p : h_{k_i}(\vec{X}) \leq z_{k_i}\} \cap B_{\underline{k}_i, \bar{k}_i}(\vec{X})$$

$i=2, \dots, n$

$$B_{k_i, \bar{k}_i}(\vec{X}) := \{\vec{X} \in \mathbb{R}_p : h_{k_i}(\vec{X}) > z_{k_i}\} \cap B_{\underline{k}_i, \bar{k}_i}(\vec{X})$$

Um einzusehen, daß die Menge $B_{\underline{k}_i, k_i}(\vec{X})$ $k_i - \underline{k}_i - 1$ Ele-

mente der verbleibenden Stichprobe enthält, verifiziere man dies für den Fall $i=1$ und $\underline{k}_1 = 0$:

$B_{0,k_1}(\vec{X})$ enthält $k_1 - 1$ Elemente. Für $i > 0$ nehme man an,

die Aussage ist richtig für alle $j < i$. Die bis zum i -ten Schritt schon "verwendeten" Indices $k_1, \dots, k_{i-1}$ ordne man der Größe nach und bezeichne diese geordneten Größen mit $l_1 < \dots < l_{i-1}$; die schon entstandenen disjunkten Mengen sind dann mit

$B_{0,l_1}(\vec{X}), B_{l_1,l_2}(\vec{X}), \dots, B_{l_{i-1},n+1}(\vec{X})$ bezeichnet.

Diese Mengen bilden eine (zufällige) Partition von $\mathbb{R}_p$.

Für k_i gilt dann:

$$k_i, \bar{k}_i \in \{0, l_1, \dots, l_{i-1}, n+1\}$$

$$B_{\underline{k}_i, \bar{k}_i}(\vec{X}) \in \{B_{0,l_1}(\vec{X}), \dots, B_{l_{i-1},n+1}(\vec{X})\};$$

nach Voraussetzung enthält die Menge $B_{\underline{k}_i, \bar{k}_i}(\vec{X})$

$\bar{k}_i - k_i - 1$ Elemente der ursprünglichen Stichprobe und der Definition des Verfahrens zur Folge ist z_{k_i} die

$(k_i - \underline{k}_i)$ -te Orderstatistik von $h_{k_i}(\vec{X}_{\frac{k_i}{k_i-1}}^{(i)}), \dots, h_{k_i}(\vec{X}_{\frac{k_i}{k_i-1}}^{(i)})$

$\vec{X}_{k_i+1}^{(i)}, \dots, \vec{X}_{k_i-1}^{(i)}$ sind hierbei die in $B_{k_i, \bar{k}_i}(\vec{X})$ liegenden Stichprobenelemente; somit verbleiben in $B_{k_i, k_i}(\vec{X})$ $k_i - k_i - 1$ Elemente und in $B_{k_i, \bar{k}_i}(\vec{X})$ $\bar{k}_i - k_i - 1$ Elemente der Stichprobe.

Das Verfahren bricht ab, wenn alle n Elemente der Stichprobe verbraucht sind, also nach n Schritten. Genauso induktiv lassen sich jetzt die Zufallsvariablen $U_{k_1}, \dots, U_{k_n}$ definieren:

$$V_{k_1} := \mu(B_{0, k_1}(\vec{X}))$$

$$U_{k_1} := V_{k_1}$$

$$V_{k_i} := \mu(B_{k_i, k_i}(\vec{X}))$$

$$U_{k_i} := \sum_{\substack{j \leq k_i \\ j \leq i}} V_{k_j}.$$

Zur Herleitung der Verteilung von $(U_{k_1}, \dots, U_{k_n})$:

Wie schon gezeigt wurde, besitzt U_{k_1} die Dichte

$$f_{k_1}(u_{k_1}) = \frac{n!}{(k_1-1)!(n-k_1)!} \cdot u_{k_1}^{k_1-1} (1-u_{k_1})^{n-k_1} I_{[0;1]}(u_{k_1});$$

nimmt man weiter an, die Dichte der gemeinsamen Verteilung von $(U_{k_1}, \dots, U_{k_j})$, $j < i$, sei bereits ermittelt und habe die Gestalt

$$f((u_{k_1}, \dots, u_{k_j})) = \frac{n!}{(l_1-1)!(l_2-l_1-1)!\dots(n-l_j)!} \cdot u_{k_1}^{l_1-1} (u_{k_2}-u_{k_1})^{l_2-l_1-1} \dots (1-u_{k_j})^{n-l_j} \\ \text{für } 0 \leq u_{k_1} \leq \dots \leq u_{k_j} \leq 1;$$

wie oben bezeichne $l_1 < l_2 < \dots < l_j$ wieder die geordneten Indices $k_1, \dots, k_j$. Für $j=2$ wurde diese Dichte schon hergeleitet. Weiter betrachte man die bedingte Dichte von V_{k_i} unter der Bedingung

$$(U_{k_1}, \dots, U_{k_{i-1}}) = (u_{k_1}, \dots, u_{k_{i-1}}), \text{ hierbei sind die}$$

$\bar{k}_i - k_i - 1$ Elemente der Stichprobe, die in die Menge $B_{\bar{k}_i, k_i}(\vec{x})$ - jetzt nicht zufällig - fallen, unabhängig. v_{k_i} ist somit die $(k_i - \bar{k}_i)$ -te Orderstatistik von $\bar{k}_i - k_i - 1$ auf $[0; v_{\bar{k}_i}]$ gleichverteilten Zufallsvariablen, die die folgende Dichte besitzt:

$$g_{k_i}(v_{k_i}) = \frac{(\bar{k}_i - k_i - 1)!}{(k_i - \bar{k}_i - 1)! (\bar{k}_i - k_i - 1)!} \cdot \frac{v_{k_i}^{k_i - \bar{k}_i - 1} (v_{\bar{k}_i} - v_{k_i})^{\bar{k}_i - k_i - 1}}{v_{\bar{k}_i}^{\bar{k}_i - k_i - 1}} \cdot I_{[0; v_{\bar{k}_i}]}(v_{k_i}) ;$$

beachtet man weiter, daß

$$v_{k_i} = u_{k_i} - u_{\bar{k}_i} \quad \text{und}$$

$$v_{\bar{k}_i} - v_{k_i} = u_{\bar{k}_i} - u_{k_i} \quad \text{gilt,}$$

so ergibt sich für die bedingte Dichte von U_{k_i} unter der Bedingung $(u_{k_1}, \dots, u_{k_{i-1}})$:

$$f_{k_i}(u_{k_i}) = \frac{(\bar{k}_i - k_i - 1)!}{(k_i - \bar{k}_i - 1)! (\bar{k}_i - k_i - 1)!} \cdot \frac{(u_{k_i} - u_{\bar{k}_i})^{k_i - \bar{k}_i - 1} (u_{\bar{k}_i} - u_{k_i})^{\bar{k}_i - k_i - 1}}{(u_{\bar{k}_i} - u_{k_i})^{\bar{k}_i - k_i - 1}}$$

wobei $u_{\bar{k}_i} \leq u_{k_i} \leq u_{\bar{k}_i}$ gilt.

Für die gemeinsame Dichte von $(U_{k_1}, \dots, U_{k_i})$ ergibt sich nun, wenn man die Menge

$\{0, 1, \dots, l_{i-1}\} \cup \{k_i\}$ der Größe nach zu $0 < l_1 < \dots < l_i < n+1$ ordnet:

$$f_{k_1}(u_{k_1}) f((u_{k_1}, \dots, u_{k_1-1})) =$$

$$\frac{n!}{(l_1-1)!(l_2-l_1-1)!\dots(n-l_1)!} \cdot$$

$$u_{l_1}^{l_1-1} \cdot (u_{l_2}-u_{l_1})^{l_2-l_1-1} \dots (1-u_{l_1})^{n-l_1}$$

$$\text{für } 0 \leq u_{l_1} \leq u_{l_2} \leq \dots \leq u_{l_1} \leq 1.$$

Nach n Schritten ergibt sich für die geordnete Stichprobe $U_1 \leq \dots \leq U_n$ von $U_{k_1}, \dots, U_{k_n}$:

$$f((u_1, \dots, u_n)) = n! \quad \text{falls } 0 \leq u_1 \leq \dots \leq u_n \leq 1 \text{ gilt.}$$

Dies ist die Dichte der Orderstatistik n unabhängiger, auf dem Intervall $[0;1]$ gleichverteilter Zufallsvariabler.

Nach n Schritten wurde gleichzeitig der Zufallsvektor $\vec{V} := (V_1, \dots, V_n)$ erzeugt,

$$V_i = \mu(B_{i-1,i}(\vec{X})), \quad i=1, \dots, n+1.$$

Für diese Zufallsgröße gilt:

$$0 \leq V_i, \quad i=1, \dots, n \quad \text{und}$$

$$\sum_{i=1}^n V_i \leq 1. \quad \text{Im weiteren sei}$$

$$D_+ := \{ \vec{V}_1 \in \mathbb{R} : 0 \leq v_1, \quad \sum_{i=1}^n v_i \leq 1 \}.$$

Der Zufallsvektor $\vec{V}$ ist *Dirichlet*-verteilt und besitzt die Dichte¹⁾:

$$f(\vec{V}) = \begin{cases} n! & \text{falls } \vec{V} \in D_+ \\ 0 & \text{sonst} \end{cases}.$$

Diese Zufallsgröße wird vor allem in §4 dieser Arbeit Verwendung finden.

1) siehe S. 25

Der Stichprobenraum $\mathbb{R}_p$ wurde nach n Schritten in $n+1$ disjunkte Teilmengen

$$B_{0,1}(\vec{x}), B_{1,2}(\vec{x}), \dots, B_{n,n+1}(\vec{x})$$

aufgeteilt. Sie seien der kürzeren Schreibweise wegen

$$B_i(\vec{x}) := B_{i-1,i}(\vec{x}), \quad i=1, \dots, n+1 \quad \text{gesetzt.}$$

Betrachtet man diese Mengen nach der Realisation der Stichprobe, so werden sie, wenn keine Verwechslung zu befürchten ist, mit $B_i(\vec{x})$, $i=1, \dots, n+1$

bezeichnet. Die Menge $\{B_1(\vec{x}), \dots, B_{n+1}(\vec{x})\}$ bildet

eine zufällige Partition von $\mathbb{R}_p$ und die Menge $\{B_1(\vec{x}), \dots, B_{n+1}(\vec{x})\}$ deren nichtzufälliges Gegenstück. Wie sich diese beiden Objekte zueinander verhalten, wird im folgenden Paragraphen untersucht werden.

Das nachfolgende Beispiel soll das in diesem Abschnitt vorgestellte Verfahren veranschaulichen.

Beispiel:

$\vec{x}_1, \dots, \vec{x}_5$ seien fünf Realisationen von unabhängigen und identisch verteilten zweidimensionalen Zufallsvektoren und wie in nachstehender Skizze angeordnet. Zur Bildung der StEB's sei die Permutation

$$\pi: \{1, \dots, 5\} \longrightarrow \{1, \dots, 5\} \quad \text{mit}$$

$$\pi(1) = k_1 = 3, \quad \pi(2) = k_2 = 2, \quad \pi(3) = k_3 = 5$$

$$\pi(4) = k_4 = 1, \quad \pi(5) = k_5 = 4$$

vorgegeben. Für die Schnittfunktionen gelte

$$h_{k_1} = h_{k_4} = h_{k_5} = pr_1 \quad \text{und}$$

$$h_{k_2} = h_{k_3} = pr_2.$$

Dies ergibt:

$$z_{k_1} = pr_1(\tilde{x}_1) , z_{k_2} = pr_2(\tilde{x}_3) , z_{k_3} = pr_2(\tilde{x}_1)$$

$$z_{k_4} = pr_1(\tilde{x}_4) , z_{k_5} = pr_1(\tilde{x}_5) .$$

Die sukzessive entstehenden Mengen sind dann

$$B_{0,3}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 : pr_1(\tilde{x}) \leq pr_1(\tilde{x}_2)\}$$

$$B_{3,6}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 : pr_1(\tilde{x}) > pr_1(\tilde{x}_2)\}$$

$$B_{0,2}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 : pr_1(\tilde{x}) \leq pr_1(\tilde{x}_2) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_3)\}$$

$$B_3(\tilde{x}) := B_{2,3}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}) \leq pr_1(\tilde{x}_2) , pr_2(\tilde{x}) > pr_2(\tilde{x}_3)\}$$

$$B_{3,5}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 : pr_1(\tilde{x}) > pr_1(\tilde{x}_2) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_1)\}$$

$$B_6(\tilde{x}) := B_{5,6}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}) > pr_1(\tilde{x}_2) , pr_2(\tilde{x}) > pr_2(\tilde{x}_1)\}$$

$$B_1(\tilde{x}) := B_{0,1}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}) \leq pr_1(\tilde{x}_4) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_3)\}$$

$$B_2(\tilde{x}) := B_{1,2}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}_4) < pr_1(\tilde{x}) \leq pr_1(\tilde{x}_2) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_3)\}$$

$$B_4(\tilde{x}) := B_{3,4}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}_2) < pr_1(\tilde{x}) \leq pr_1(\tilde{x}_5) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_1)\}$$

$$B_5(\tilde{x}) := B_{4,5}(\tilde{x}) = \{\tilde{x} \in \mathbb{R}_2 :$$

$$pr_1(\tilde{x}) > pr_1(\tilde{x}_5) , pr_2(\tilde{x}) \leq pr_2(\tilde{x}_1)\}$$

Skizze:



1.4 Gegenstand der Arbeit

Ausgangspunkt der Arbeit war der Beitrag T. W. Anderson's zu einem internationalen Symposium in Dayton/Ohio im Jahre 1965 über multivariate Analysis, in dem er sich mit Statistically Equivalent Blocks beschäftigte;¹⁾ vor allem war seine Aussage über die Konsistenz des Verfahrens: Die StEB's "konvergieren" (stochastisch) in dem von ihm angeführten Beispiel gegen vorgegebene Partitionen, nicht so leicht zu beweisen, wie es zunächst den Anschein hatte. So beschäftigt sich Paragraph 2 damit, zunächst einen geometrischen Zugang zu dieser Aussage in einer allgemeineren Form zu finden. Im einzelnen werden notwendige und hinreichende Bedingungen dafür angegeben, wann eine Folge von (nicht zufallsbedingten) Partitionen ein Wahrscheinlichkeitsmaß eindeutig festlegt; weiter wird definiert, was unter einer konvergenten Folge von Partitionen und unter einer Folge von StEB's, die gegen eine Folge von Partitionen schwach konvergieren, zu verstehen ist. Es konnten mit den Mitteln dieses Abschnitts keine Bedingungen aufgezeigt werden, unter denen eine Folge von StEB's gegen eine Folge von Partitionen konvergiert. Eine Beantwortung dieser Frage führte zu der in Paragraph 3 entwickelten Betrachtungsweise.

In diesem Abschnitt wird ein auf Abbildungen beruhender Zugang zur Beantwortung der oben gestellten Frage entwickelt, der es darüber hinaus möglich macht, unter die Allgemeinheit nur geringfügig

1) Siehe Seite 7 unter ²⁾

gig einschränkenden Annahmen an den zugrunde liegenden Hypothesenraum zu zeigen, wann die Verteilung der zugrundeliegenden Stichprobe vom Umfang n in eindeutiger Weise durch das Verfahren in eine Verteilung auf dem Meßraum $(\mathbb{R}_n; \mathcal{B}_n)$ "transformiert" werden kann. Am Ende dieses Abschnitts wird noch die von T. W. Anderson behauptete "Konsistenz" mittels der in §2 und §3 bereitgestellten Mittel gezeigt. In §4 wird, auf eine in der Literatur schon bekannten Testgröße zurückgreifend,¹⁾ ein Test entwickelt, wobei bei der Wahl der Testgröße deren Unabhängigkeit von der Reihenfolge der Bildung der StEB's ausschlaggebend war. Auch bietet sie eine Möglichkeit Gütebetrachtungen anzustellen; auf Betrachtungen dieser Art wird in der vorliegenden Arbeit nicht eingegangen.

Die abschließenden Bemerkungen des Paragraphen 5 beinhalten Fragen, die in dieser Arbeit offen blieben, nicht behandelt wurden oder erst durch sie entstanden sind.

Ein Literaturverzeichnis befindet sich am Ende der Arbeit.

1) Sherman, B.: A random variable related to the Spacing of sample values. Ann. Math. Statistics 21, 339 - 361 (1950).

§ 2

Konsistenz und Statistically Equivalent Blocks

Betrachtet man die Vorgehensweise StEB's zu bilden, so stellt sich sofort die Frage:

Inwieweit und unter welchen (einschränkenden) Bedingungen ist es dem Verfahren möglich, zu erkennen, wenn eine andere Verteilung μ_1 als diejenige (μ_0), mit der die StEB's integriert wurden, zugrunde liegt? Welche Forderungen sind ggf. an die in Betrachtung gezogenen Verteilungen (Hypothesenraum) zu stellen, um diese Unterscheidbarkeit zu gewährleisten?

Einen ersten Schritt, sich der Beantwortung dieser Fragen zu nähern, tut T. W. Anderson¹⁾, indem er sich die Frage nach der Konsistenz des Verfahrens stellt. Seine Vorgehensweise sei im Folgenden kurz wiedergegeben:

Zugrunde gelegt ist eine übliche Definition konsistenter Testfolgen:

Sei $((\phi_n; \alpha_n): n \in \mathbb{N})$ eine Testfolge zu den Niveaus $\alpha_n \in (0; 1)$ und $(X_1, \dots, X_n) := \vec{X}_n - n$ bezeichne den Stichprobenumfang und μ_0 die Hypothese -, dann heißt $((\phi_n; \alpha_n): n \in \mathbb{N})$ konsistent gegen eine Alternative $\mu_1 \neq \mu_0$, wenn gilt:

$$\lim_{n \rightarrow \infty} W(\phi_n(\vec{X}_n) = 1 | \mu_1) = 1.$$

¹⁾ Siehe Seite 7 unter 2)

Weiter geht T. W. Anderson davon aus, daß der Test, der auf die durch StEB's entstandenen Zufallsvariablen angewendet wird, konsistent ist, wie z.B. der Kolmogorov- und Cramér - von Mises-Test. Seine Argumentation gründet sich dann auf folgendes Vorgehen:

Der Stichprobenraum $\mathbb{R}$ wird in zwei lineare Halbräume

$$B(0, \frac{1}{2}) := \{\vec{x} \in \mathbb{R}_2 : pr_1(\vec{x}) \leq a_1\} \quad \text{und}$$

$$B(\frac{1}{2}, 1) := \{\vec{x} \in \mathbb{R}_2 : pr_1(\vec{x}) > a_1\} \quad \text{eingeteilt,}$$

wobei $a_1 \in \mathbb{R}$ so gewählt ist, daß

$$\mu_0(B(0, \frac{1}{2})) = \mu_0(B(\frac{1}{2}, 1)) = \frac{1}{2} \quad \text{gilt.}$$

In den beiden folgenden Schritten werden diese beiden Mengen jeweils ebenfalls in zwei Mengen aufgeteilt.

$$B(0, \frac{1}{4}) := B(0, \frac{1}{2}) \cap \{\vec{x} \in \mathbb{R}_2 : pr_2(\vec{x}) \leq a_2\}$$

$$B(\frac{1}{4}, \frac{1}{2}) := B(0, \frac{1}{2}) \cap \{\vec{x} \in \mathbb{R}_2 : pr_2(\vec{x}) > a_2\}$$

$$B(\frac{1}{2}, \frac{3}{4}) := B(\frac{1}{2}, 1) \cap \{\vec{x} \in \mathbb{R}_2 : pr_2(\vec{x}) \leq a_3\}$$

$$B(\frac{3}{4}, 1) := B(\frac{1}{2}, 1) \cap \{\vec{x} \in \mathbb{R}_2 : pr_2(\vec{x}) > a_3\};$$

hierbei sind $a_2 \in \mathbb{R}$ und $a_3 \in \mathbb{R}$ so gewählt, daß

$$\mu_0(B(0, \frac{1}{4})) = \mu_0(B(\frac{1}{4}, \frac{1}{2})) = \mu_0(B(\frac{1}{2}, \frac{3}{4})) =$$

$$= \mu_0(B(\frac{3}{4}, 1)) = \frac{1}{4} \quad \text{gilt.}$$

T. W. Andersons Aussage läßt sich, die Spezialisierung auf den zweidimensionalen Fall außer acht lassend, wie folgt zusammenfassen:

Teilt man den Stichprobenraum sukzessive in Teilmengen ein, die die Verteilung der Hypothese - μ_0 - eindeutig festlegen, und ist es weiter möglich, die StEB's so zu konstruieren, daß sie diese Mengen bzgl. μ_0 beliebig gut approximieren, so ist, vorausgesetzt der darauf aufbauende Test ist konsistent, das ganze Verfahren konsistent. Diese Argumentationskette zeigt die drei Punkte, die in Hinblick auf die Konsistenz des Verfahrens zu klären sind:

- Wann ist ein Wahrscheinlichkeitsmaß μ_0 durch Werte auf einer Folge von Mengen (die speziell für dieses Wahrscheinlichkeitsmaß gewählt sind) eindeutig von anderen Wahrscheinlichkeitsmaßen unterscheidbar?
- Was bedeutet die Aussage: 'die StEB's approximieren eine vorgegebene Folge von Mengen mit wachsendem Stichprobenumfang immer besser'?
- Darauf aufbauend stellen sich die Fragen:

Ist für ein Wahrscheinlichkeitsmaß μ_0 dann eine Folge von Mengen angebbar, so daß μ_0 von anderen Wahrscheinlichkeitsmaßen durch seine Werte auf dieser Mengenfolge unterscheidbar ist?

Kann zu dieser Folge eine Konstruktionsweise von StEB's gefunden werden, so daß diese StEB's die obige Folge immer besser approximieren?

Den beiden ersten Fragen systematisch nachzugehen, ist Gegenstand dieses Paragraphen; Die Behandlung des dritten Punktes ist erst mittels der in § 3 entwickelten Abbildungen möglich. Dieser Paragraph soll zum einen eine anschauliche geometrische Motivation der im folgenden Paragraphen zu definierenden Abbildungen vermitteln und zum anderen die Grenzen der geometrischen Betrachtungsweise aufzeigen. Weiter beinhaltet dieser Paragraph die Rechtfertigung, im weiteren Verlauf der Arbeit nur noch Projektionen auf die Koordinatenachsen als Schnittfunktionen zu betrachten.

2.1 Partitionenfolgen, die ein Wahrscheinlichkeitsmaß eindeutig bestimmen

Einführend einige Definitionen aus der Maßtheorie und Mengenlehre:

- Unter der *Partition* einer Menge E versteht man eine endliche oder unendliche Klasse P von disjunkten nichtleeren Teilmengen E' der Menge E , für die gilt:

$$\bigcup_{E' \in P} E' = E.$$
- Sind P_1 und P_2 Partitionen der gleichen Menge E , so heißt P_1 *feiner als* P_2 ($P_1 \leq P_2$), wenn jedes Element von P_1 Teilmenge eines Elementes von P_2 ist.
- Eine *Folge von Partitionen* $(P_n : n \in \mathbb{N})$ heißt *fallend*, wenn gilt: $P_{i+1} \leq P_i, \quad i \in \mathbb{N}.$
- Sei $(\Omega; \mathcal{A}_0)$ ein Maßraum und μ ein Maß auf $(\Omega; \mathcal{A}_0)$, so heißt $N \in \mathcal{A}_0$ μ - Nullmenge, wenn $\mu(N) = 0$ gilt.
- A und B seien Mengen;

$$A \Delta B := (A \cap \bar{B}) \cup (B \cap \bar{A}) \quad (\text{symmetrisches Produkt}).$$

- Existiert zu einem Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}_p; \mathcal{B}_p)$ eine positive Riemann - integrierbare Funktion f_μ , so daß für alle $A \in \mathcal{B}_p$

$$\mu(A) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_\mu((x_1, \dots, x_p)) I_A((x_1, \dots, x_p)) dx_p \dots dx_1$$

gilt, so heie μ stetig und f_μ Dichte von μ .

Ausgegangen wird in diesem Paragraphen, wenn nicht ausdrcklich etwas anderes vermerkt wird, von $(\mathbb{R}_p; \mathcal{B}_p)$, $p \geq 2$ und stetigen Wahrscheinlichkeitsmaen auf diesem Meraum. $A(E)$ bezeichne immer die kleinste Algebra und $\mathcal{A}_\sigma(E)$ die kleinste σ -Algebra ber der Menge E . Die folgende Definition diene einer einfacheren und zusammenfassenderen Sprechweise.

2.1.1 Definition:

Sei μ ein (stetiges) Wahrscheinlichkeitsma auf $(\mathbb{R}_p; \mathcal{B}_p)$ und $(P_n: n \in \mathbb{N})$ eine Folge von Partitionen von $\mathbb{R}_p$, dann heie $(P_n: n \in \mathbb{N})$ eine (μ) -Partitionenfolge, wenn gilt:

Ist ν ein (stetiges) Wahrscheinlichkeitsma auf $(\mathbb{R}_p; \mathcal{B}_p)$ und $\mu(B_n) = \nu(B_n)$ fr alle $n \in \mathbb{N}$ und alle $B_n \in P_n$, so folgt $\nu = \mu$.

Ein Wahrscheinlichkeitsma μ ist durch seine Werte auf einer solchen Partitionenfolge eindeutig bestimmt. Eine Charakterisierung von (μ) -Partitionenfolgen ist Gegenstand des nachfolgenden Satzes.

2.1.2 Satz:

Eine absteigende Folge von Partitionen $(P_n: n \in \mathbb{N}) \in \mathcal{P}_P^{\mathbb{N}}$ auf $\mathbb{R}_P$ ist genau dann eine (μ) -Partitionenfolge für ein stetiges Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}_P; \mathcal{B}_P)$, wenn für alle $A \in \mathcal{B}_P$ gilt:

$$\begin{aligned} (+) \quad \mu^*(A) &:= \inf_{n \in \mathbb{N}} \{ \mu(A \cap B_n) : A \subset B_n \in \mathcal{A}(P_n), n \in \mathbb{N} \} = \\ \mu_*(A) &:= \sup_{n \in \mathbb{N}} \{ \mu(A \cup C_n) : A \supset C_n \in \mathcal{A}(P_n), n \in \mathbb{N} \}. \end{aligned}$$

Beweis:

$\Leftarrow$: es ist zu zeigen: gilt für alle $A \in \mathcal{B}_P$ die Gleichung (+), so ist $(P_n: n \in \mathbb{N})$ eine (μ) -Partitionenfolge.

Hierzu sei ν ein weiteres Wahrscheinlichkeitsmaß auf $(\mathbb{R}_P; \mathcal{B}_P)$ und

$$\begin{aligned} A_\sigma((P_n: n \in \mathbb{N}))^\mu &= (A \Delta N: A \in A_\sigma((P_n: n \in \mathbb{N})), N \subset N' \in A_\sigma((P_n: n \in \mathbb{N}))) \\ &\text{und } \mu(N') = 0. \end{aligned}$$

Hiermit gilt dann:

$$\mu(B_n) = \nu(B_n) \quad \text{für alle } n \in \mathbb{N} \text{ und für alle } B_n \in P_n$$

$$\Rightarrow \mu = \nu \quad \text{auf } \mathcal{A}((P_n: n \in \mathbb{N}))$$

$$\Rightarrow \mu = \nu \quad \text{auf } A_\sigma((P_n: n \in \mathbb{N}))$$

$$\Rightarrow \mu = \nu \quad \text{auf } \overline{A_\sigma((P_n: n \in \mathbb{N}))}^\nu = \overline{A_\sigma((P_n: n \in \mathbb{N}))}^\mu,$$

Aus (+) folgt dann

$$A_\sigma((P_n: n \in \mathbb{N})) \subset \mathcal{B}_P \subset \overline{A_\sigma((P_n: n \in \mathbb{N}))}^\mu \quad \text{d.h.}$$

für alle $A \in \mathcal{B}_P$ gilt $A = A_N \Delta N$ wobei

$A_N \in A_\sigma((P_n: n \in \mathbb{N}))$ und N Teilmenge einer

μ - bzw. ν - Nullmenge ist.

$\bar{\mu}$ und $\bar{\nu}$ bezeichnen jeweils die Kompletterweiterungen der Wahrscheinlichkeitsmaße μ und ν auf $\overline{A_\sigma((P_n: n \in \mathbb{N}))}$;

Somit gilt für $A \in \mathcal{E}_P$ und $A = A_N \Delta N$

$$\bar{\mu}(A_N \Delta N) = \mu(A_N) = \mu(A_N \Delta N') = \mu(A) \quad \text{und}$$

$$\bar{\nu}(A_N \Delta N) = \nu(A_N) = \nu(A_N \Delta N') = \nu(A) \quad \text{daraus folgt}$$

$$\mu(A) = \nu(A).$$

" $\implies$ ":

Es soll bewiesen werden: ist eine Partitionenfolge eine (μ) -Partitionenfolge, so genügt sie der Gleichung (+). Dies wird durch Widerspruch gezeigt.

Annahme:

Es existiert ein $A \in \mathcal{E}_P$, für das ohne Einschränkung

$$\mu_*(A) < \mu(A) = \mu^*(A) \text{ gilt.}$$

Dann gibt es ein $Q \in \mathcal{E}_P \setminus \mathcal{A}((\mathcal{F}_n : n \in \mathbb{N}))$ mit den Eigenschaften:

für $n \in \mathbb{N}$ und $B_n \in \mathcal{F}_n$ gilt

$$Q \cap B_n \neq \emptyset \implies Q \subset B_n,$$

$$\mu(Q) = \alpha > 0,$$

$$\lambda_P(Q) = \beta < \infty.$$

f_μ sei die Dichte des Wahrscheinlichkeitsmaßes μ , weiter sei $Q_1 \subset Q$, $Q_2 \subset Q$, $Q_1 \cap Q_2 \neq \emptyset$ und $Q_1 \cup Q_2 = Q$.

o.E. kann angenommen werden, daß Q , Q_1 und Q_2 beschränkte Quader in $\mathbb{R}_P$ sind und

$$\mu(Q_1) = \alpha_1 > 0 \text{ sowie } \mu(Q_2) = \alpha_2 > 0 \text{ gilt.}$$

$$\lambda_P(Q_1) := \beta_1 > 0$$

$$\lambda_P(Q_2) := \beta_2 > 0$$

Hiermit läßt sich nun die folgende Funktion definieren:

$$f_{\mu}^1(\vec{x}) := \begin{cases} f_{\mu}(\vec{x}) & \text{für } \vec{x} \in \bar{Q} \\ \frac{\gamma \cdot \alpha}{\beta_1} & \text{für } \vec{x} \in Q_1 \\ \frac{(1-\gamma) \cdot \alpha}{\beta_2} & \text{für } \vec{x} \in Q_2 \end{cases}$$

$$\gamma \in (0; \frac{\alpha_1}{\alpha})$$

f_{μ}^1 ist eine meßbare, reelle Abbildung und für $\vec{x} \in \mathbb{R}_p$ ist $f_{\mu}^1(\vec{x}) \geq 0$.

Weiter gilt:

$$\int_{\mathbb{R}_p} f_{\mu}^1(\vec{x}) d\vec{x} = \int_{\bar{Q}} f_{\mu}(\vec{x}) d\vec{x} + \int_{Q_1} \frac{\gamma \cdot \alpha}{\beta_1} d\vec{x} + \int_{Q_2} \frac{(1-\gamma) \alpha}{\beta_2} d\vec{x} =$$

$$\mu(\bar{Q}) + \gamma \cdot \alpha + (1-\gamma) \alpha = \mu(\bar{Q}) + \alpha = 1.$$

Somit ist f_{μ}^1 eine Dichte und μ_1 bezeichne das dazugehörige Wahrscheinlichkeitsmaß.

Es gilt:

$$\mu_1(Q_1) = \frac{\gamma \cdot \alpha}{\beta_1} \quad \lambda_p(Q_1) = \gamma \cdot \alpha < \alpha_1 = \mu(Q_1) \implies \mu \neq \mu_1$$

Dies ist jedoch ein Widerspruch zur Annahme, da für

$$B_n \in \mathcal{P}_n, \quad n \in \mathbb{N}$$

$$\mu_1(B_n) := \mu(B_n) \quad \text{falls } B_n \cap Q = \emptyset \quad \text{sowie}$$

$$\mu_1(B_n) = \mu_1(B_n \setminus Q) + \mu_1(Q) = \mu(B_n \setminus Q) + \alpha =$$

$$\mu(B_n) \quad \text{gilt.}$$

//.

Würde man die Aussage des obigen Satzes etwas modifizieren und μ_* und μ^* direkt als inneres und äußeres Maß formulieren, so könnte er mit Hilfe eines Satzes aus der Maßtheorie¹⁾ bewiesen werden.

Dieser Satz besagt, daß auf der Vervollständigung einer σ - Algebra bzgl. eines Maßes μ die Abbildungen μ^* und μ_* identisch sind und umgekehrt, die Mengen E eines "hereditary - σ - Rings" über einer Algebra A , für die

$$\mu^*(E) = \mu_*(E) < \infty$$

gilt, bereits zur Vervollständigung der von A erzeugten σ - Algebra bzgl. des Maßes μ gehören.

Um den Beweis mit Hilfe dieser Aussagen zu führen, wäre es nötig gewesen, zunächst den Begriff "hereditary - σ - Ring" zu erläutern und dann zu beweisen, daß

$$B_p \subset \overline{A_\sigma((P_n : n \in \mathbb{N}))}^*$$

ist.

Da vor allem der Nachweis der letzten Aussage technisch und praktisch der gleiche wäre, wie der im Beweis des Satzes geführte, wurde der direkte Weg gewählt.

1) HALMOS P.R.: Measure Theory, D. van Nostrand Company Inc. Princetown, New Jersey, Toronto, New York, London (1950)

Es ist zu beachten, daß eine (μ) -Partitionenfolge nicht notwendig andere Wahrscheinlichkeitsmaße festlegt; es kann z.B. für zwei weitere Wahrscheinlichkeitsmaße $\rho \neq \mu$, $\nu \neq \mu$ gelten:

$\rho(B_{j_n}) = \nu(B_{j_n})$, $n \in \mathbb{N}$ und $B_{j_n} \in \mathcal{F}_n$
obwohl $\rho \neq \nu$ ist.

Bei der Konstruktion von StEB's werden Abbildungen $h_1, \dots, h_m$, $m \leq n$, $h_i \neq h_j$, $i \neq j$, die Schnittfunktionen, verwendet. Man betrachte nun eine Folge von Realisationen einer unabhängigen Stichprobe.

Dies ist eine Folge $(\tilde{x}_n; n \in \mathbb{N}) \in \mathbb{R}_p^{\mathbb{N}}$. Die ersten k Glieder dieser Folge ermöglichen es, $k + 1$ StEB's zu bilden, somit definiert $(\tilde{x}_n; n \in \mathbb{N})$ eine Folge von Partitionen. Weiter ist jedes Element einer Partition Urbild eines geeigneten, nicht notwendig beschränkten Quaders aus $\mathbb{R}_m$ bezüglich der Abbildung $H := (h_1, \dots, h_m) : \mathbb{R}_p \longrightarrow \mathbb{R}_m$.

Dies sei am folgenden Beispiel gezeigt:

Sei $H := (h_1, h_2, h_3)$, $n = 4$,

$k_1 = 4$, $k_2 = 3$, $k_3 = 1$, $k_4 = 2$,

$h_{k_1} = h_{k_3} = h_1$, $h_{k_2} = h_2$, $h_{k_4} = h_3$

z_{k_1} , z_{k_2} , z_{k_3} , z_{k_4} seien aus der Stichprobe ermittelt

und es gelte $z_{k_1} > z_{k_3}$. Damit sind die folgenden

Mengen zu konstruieren:

$$B_{0,4} \quad , \quad B_{4,5} \quad ,$$

$$B_{0,3} \quad , \quad B_{3,4} \quad ,$$

$$B_{0,1} \quad , \quad B_{1,3} \quad ,$$

$$B_{1,2} \quad , \quad B_{2,3} \quad , \text{ der einfachen Schreibweise}$$

wegen wurde auf " $(\vec{x})$ " verzichtet.

Betrachtet man nun die Menge $B_{2,3}$, so schreibt sie sich folgendermaßen

$$\underbrace{\{\vec{x} \in \mathbb{R}_p : h_1(\vec{x}) \leq z_{k_1}\}}_{B_{0,4}} \cap \underbrace{\{\vec{x} \in \mathbb{R}_p : h_2(\vec{x}) \leq z_{k_2}\}}_{B_{0,3}} \cap \underbrace{\{\vec{x} \in \mathbb{R}_p : h_1(\vec{x}) > z_{k_3}\}}_{B_{1,3}}$$

$$B_{2,3} = B_{1,3} \cap \{\vec{x} \in \mathbb{R}_p : h_3(\vec{x}) > z_{k_4}\} =$$

$$h_1^{-1}((z_{k_2}; z_{k_1}]) \cap h_2^{-1}((-\infty; z_{k_2}]) \cap h_3^{-1}((z_{k_4}; \infty)) =$$

$$H^{-1}(\text{pr}_1^{-1}((z_{k_3}; z_{k_1}]) \cap \text{pr}_2^{-1}((-\infty; z_{k_2}]) \cap$$

$$\cap \text{pr}_3^{-1}((z_{k_4}; \infty))) = H^{-1}(Q)$$

$$\text{wobei } Q = \{(x_1, x_2, x_3) \in \mathbb{R}_3 : z_{k_3} < x_1 \leq z_{k_1}, -\infty < x_2 \leq z_{k_2},$$

$$z_{k_4} < x_3 < \infty\} \text{ gilt.}$$

Betrachtet man jetzt nicht mehr StEB's, sondern "gewöhnliche" Folgen von Partitionen, so kann man, in Anlehnung an die obigen Überlegungen, solche Partitionen betrachten, die durch eine Abbildung $H := (h_1, \dots, h_m)$ aus einer Folge von Partitionen aus $\mathbb{R}_m$, deren Elemente aus Quadern (nicht notwendig nach allen Richtungen beschränkt) bestehen, entstanden sind. Die reellen Abbildungen $h_1, \dots, h_m$ seien auch jetzt wieder, um die Analogie zu verdeutlichen, als Schnittfunktionen angesprochen. Die folgenden beiden Sätze formulieren Aussagen über die Funktion $H := (h_1, \dots, h_m)$, wenn die Partitionsfolge $(P_n: n \in \mathbb{N})$, die durch H aus einer aus Quadern bestehenden Partitionsfolge $(Q_n: n \in \mathbb{N})$ aus $\mathbb{R}_m$ hervorgegangen ist, eine (μ) -Partitionenfolge ist.

Die folgenden Bezeichnungsweisen und Sachverhalte liegen den weiteren Betrachtungen zugrunde.

- $\{h_1, \dots, h_m\}$ sei die Menge reeller Funktionen

$$h_i: \mathbb{R}_p \rightarrow \mathbb{R}, \quad i = 1, \dots, m$$

die als Schnittfunktionen dienen

- sei $H := (h_1, \dots, h_m)$, dann ist für alle $n \in \mathbb{N}$

und $i \in \{1, \dots, n+1\}$,

$P_n := \{B_1^n, \dots, B_{n+1}^n\}$ eine Partition. Die dazugehörige Partitionsfolge ist $(P_n: n \in \mathbb{N})$; jedes ihrer Folgenglieder enthält eine Menge mehr, als ihr Vorgänger.

darstellbar als $B_i^n = H^{-1}(Q_i^n)$, wobei $Q_i^n \subset \mathbb{R}_m$

und $Q_i^n = \bigcap_{j=1}^m \text{pr}_j^{-1}((-\infty, c_j]) \cap \bigcap_{j=1}^m \text{pr}_j^{-1}((d_j, \infty))$

für geeignete $c_j, d_j \in \mathbb{R}$, $d_j < c_j$, $j = 1, \dots, m$.

2.1.3 Satz:

Ist $(P_n: n \in \mathbb{N})$ eine für das stetige Wahrscheinlichkeitsmaß μ auf $(R_P, \mathcal{B}_P)$ absteigende (μ) -Partitionenfolge; $h_1, \dots, h_m$ seien reelle Funktionen, so daß für $n \in \mathbb{N}$ und $i = 1, \dots, n+1$ $B_i^n = H^{-1}(Q_i^n)$, wobei $H := (h_1, \dots, h_m)$ gilt und Q_i^n geeignete Quader in $\mathbb{R}_m$ sind, so gibt es eine μ -Nullmenge $N \in \mathcal{B}_P$ mit der Eigenschaft: für alle $\vec{x}, \vec{y} \in \bar{N}$ gilt $H(\vec{x}) = H(\vec{y}) \Rightarrow \vec{x} = \vec{y}$

Beweis:

Sei $P_n := \{B_1^n, \dots, B_{n+1}^n\}$, $n \in \mathbb{N}$.

Wegen der obigen Anmerkungen existieren für alle $n \in \mathbb{N}$ und alle $j=1, \dots, n+1$ Quader $Q_j^n \in \mathcal{B}_m$ mit der Eigenschaft $H^{-1}(Q_j^n) = B_j^n$ wobei für $i \neq j$ $Q_j^n \cap Q_i^n = \emptyset$ gilt; somit liegt jedes $\vec{c} \in H(\mathbb{R}_P)$ in genau einem Quader Q_j^n , $n \in \mathbb{N}$, daraus folgt $H^{-1}(\{\vec{c}\})$ ist für jedes $n \in \mathbb{N}$ in genau einem B_j^n enthalten, somit gilt, da P_n eine (μ) -Partitionenfolge ist, $\mu(H^{-1}(\{\vec{c}\})) = 0$.

- i) $Z := \{\vec{z} \in \mathbb{R}_m : |H^{-1}(\{\vec{z}\})| > 1\}$
- ii) für $\vec{z} \in Z$ sei $\bar{z}$
 $\bar{z} := \{\vec{y} \in \mathbb{R}_m : (H^{-1}(\{\vec{z}\}); H^{-1}(\{\vec{y}\})) \subset B_1^n, n \in \mathbb{N}\}$
 $\bar{y} = \bar{z}$ für alle $\vec{y} \in Z \cap \bar{z}$.
- iii) $\bar{z}$ sei die Menge aller $\bar{z}$, $z \in Z$, und enthält somit höchstens abzählbar viele Elemente;
 Weiter gilt $Z \subset \bigcup \bar{z}$
- iv) $\hat{Z}$ sei die kleinste in B_{p_n} enthaltene Menge, die $\bar{z}$ enthält, daraus folgt $\hat{Z} \subset B_1^n$ für alle $n \in \mathbb{N}$
 und somit gilt $\mu(\hat{Z}) = 0$.
- v) Insgesamt gilt:
 $H^{-1}(Z) \subset H^{-1}(\bigcup_{\vec{z} \in Z} \bar{z}) = \bigcup_{\vec{z} \in Z} H^{-1}(\bar{z}) \subset (\bigcup_{\vec{z} \in \hat{Z}} \hat{Z}) := \hat{Z} \in B_p$
 und $\mu(\bigcup_{\vec{z} \in Z} \hat{Z}) \leq \Sigma \mu(\hat{Z}) = 0 \implies$
 $H^{-1}(Z)$ ist in einer μ - Nullmenge von B_p enthalten.

Aus i) bis v) folgt:

Sind $\vec{x}, \vec{y} \in \hat{Z}$ und $H(\vec{x}) = H(\vec{y}) = \vec{z}$, so gilt:
 $H(\vec{x})$ und $H(\vec{y})$ sind nicht in Z enthalten.

$$|H^{-1}(H(\vec{x}))| = 1 \iff |H^{-1}(\vec{z})| = 1 \implies \vec{x}, \vec{y} \in H^{-1}(\vec{z}) \implies \vec{x} = \vec{y}.$$

//.

Infolge des Satzes 2.1.3 kann man sich bei der Konstruktion von (μ) -Partitionenfolgen, auf solche beschränken, die mit injektiven Abbildungen $H = (h_1, \dots, h_m)$ konstruiert werden; weiter ist auch o.E. $p = m$ zu fordern. Lemma 2.1.4 formuliert den Zusammenhang zwischen (μ) -Partitionenfolgen, die auf Projektionen aufbauen und Partitionenfolgen, die aus Urbildern bezüglich einer Abbildung H dieser Quader bestehen. Dies kann deswegen von Interesse sein, weil aus Quadern bestehende (μ) -Partitionenfolgen leicht zu konstruieren sind; ist man trotzdem an anderen Mengen interessiert, kann man diese aus Quadern mittels geeigneter Transformationen gewinnen.

2.1.4 Lemma:

Sei $H := (h_1, \dots, h_p)$ eine bimeßbare¹⁾ Abbildung des Wahrscheinlichkeitsraumes $(\mathbb{R}_p; \mathcal{B}_p; \mu)$ in den Meßraum $(\mathbb{R}_p; \mathcal{B}_p)$, so ist $(Q_n: n \in \mathbb{N})$, $Q_n := \{Q_1^n, \dots, Q_{n+1}^n\}$ genau dann eine $(H(\mu))$ -Partitionenfolge von Quadern auf $(\mathbb{R}_p; \mathcal{B}_p)$, wenn $(P_n: n \in \mathbb{N})$, $P_n := \{B_1^n, \dots, B_{n+1}^n\}$, wobei $B_i^n := H^{-1}(Q_i^n)$ für $n \in \mathbb{N}$ und $i=1, \dots, n+1$, eine (μ) -Partitionenfolge auf $(\mathbb{R}_p; \mathcal{B}_p)$ ist.

Da im weiteren Verlauf der Arbeit von der obigen Aussage kein wesentlicher Gebrauch gemacht werden wird, sei auf einen Beweis verzichtet.

Plant man einen Test (Hypothese μ_0), der mit Hilfe von StEB's durchgeführt werden soll, so ist es sinnvoll, die StEB's so zu konstruieren, daß sie, in einem noch zu präzisierenden Sinn, gegen eine (μ_0) -Partitionenfolge konvergieren, wenn μ_0 die tatsächlich zugrunde liegende Verteilung ist. Eine erste Antwort darauf, wie diese Konvergenz definiert werden kann, gibt der folgende Abschnitt.

2.2 Folgen von Partitionen und Statistically Equivalent Blocks

Dieser Abschnitt beschäftigt sich, wie bereits angekündigt, mit den Fragen 'was ist unter einer konvergenten Folge von StEB's zu verstehen?' und 'wie kann

¹⁾ Eine Abbildung H heiße bimeßbar, wenn sie bijektiv und in "beiden Richtungen" meßbar ist.

eine solche Folge eine andere Verteilung μ als die zugrunde liegende erkennen?'. Dazu wird es nötig sein, die StEB's in Abhängigkeit des Stichprobenumfanges n neu zu erklären, um dann einen vernünftigen Konvergenzbegriff formulieren zu können. Dies wird vor allem dadurch geschehen, eine Folge von StEB's als *induktives System* aufzufassen.

Wie üblich, sind μ und ν stetige Wahrscheinlichkeitsmaße auf $(\mathbb{R}_p; \mathcal{B}_p)$, $p \geq 2$; mit $\vec{X}$ seien die zufälligen Größen und mit $\vec{x}$ deren Realisationen bezeichnet.

Zunächst einige heuristische Bemerkungen:

Es könnte beispielsweise sein, daß z_{k_1} , jetzt aus der Stichprobe $\vec{X}_1, \dots, \vec{X}_{n+1}$ gewonnen, eine Koordinate des hinzugewonnenen Zufallsvektors $\vec{X}_{n+1}$ ist. Diese weitere Beobachtung $\vec{X}_{n+1}$ erfordert somit eine völlig neue "Organisation" der $\vec{X}_1, \dots, \vec{X}_{n+1}$ im Hinblick auf die Bildung der StEB's, um eine weitere Partition, die jetzt ein Element mehr als die vorangegangene besitzt, zu erhalten.

Im Anschluß einige Bezeichnungen und Festlegungen, die im weiteren dieses Paragraphen Verwendung finden

- n sei der Stichprobenumfang; $\vec{X}_1, \dots, \vec{X}_n$ seien unabhängige und identisch verteilte Zufallsvektoren, $z_{k_1, n}, \dots, z_{k_n, n}$, die durch Anwendung der Permutation und der Projektionen ermittelten "Koordinaten" und $z_{k_1, n}, \dots, z_{k_n, n}$ deren entsprechende Realisationen. Das n im Index gibt an, welcher Stichprobenumfang der Blockbildung zugrundeliegt.

- der Blockbildung liegen als Schnittfunktionen nur die p Projektionen auf die Koordinatenachsen zugrunde, sie werden jedoch der einfacheren Darstellung wegen mit $h_{k_1(n)}, \dots, h_{k_n(n)}$ bezeichnet.

- Wie in § 1 sind die Mengen

$$B_{\underline{k}_i, k_i}(\vec{X}^n), B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n), i = 1, \dots, n$$

zu verstehen.

Zusammengefaßt werden diese Mengen zu der Menge

$$P(\vec{X}^n) := \{B_{\underline{k}_i, k_i}(\vec{X}^n) : i = 1, \dots, n\} \cup \\ \{B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n) : i = 1, \dots, n\} \cup \\ \{\mathbb{R}_p, \emptyset\}$$

Betrachtet man diese Menge *nach* einer Realisation von $z_{k_1(n)}, \dots, z_{k_n(n)}$, so enthält sie eine nicht-triviale Partition von $\mathbb{R}_p$.

- Sei $(A; <)$ eine geordnete Menge und $(E_\alpha : \alpha \in A)$ eine Familie von Mengen E_α , auf denen eine Verknüpfung $v : E_\alpha \times E_\alpha \longrightarrow E_\alpha$ definiert ist; ferner sei:

$$i) \quad r_{\alpha, \beta} : E_\alpha \longrightarrow E_\beta \quad \alpha, \beta \in A \text{ und } \alpha < \beta \\ r_{\alpha, \beta} \text{ ist ein Homomorphismus.}$$

- ii) Sei $\alpha, \beta, \gamma \in A$ und $\alpha < \beta < \gamma$, so ist

$$r_{\alpha, \gamma} = r_{\beta, \gamma} \circ r_{\alpha, \beta} \\ (E_\alpha; r_{\alpha, \beta}; \alpha < \beta; \alpha, \beta \in A) \text{ heit direktes (induktives) System.}$$

- Gilt zudem:

Es existiert eine Menge E , auf der eine Verknüpfung $v : E \times E \longrightarrow E$ definiert ist und die Homomorphismen

$$r_\alpha : E_\alpha \longrightarrow E, \text{ so da } E = \bigcup_{\alpha \in A} r_\alpha(E_\alpha), \alpha \in A,$$

$$r_\alpha(E_\alpha) \subset r_\beta(E_\beta) \quad \text{fr } \alpha, \beta \in A \text{ und } \alpha < \beta$$

$$r_\beta \circ r_{\alpha, \beta} = r_\alpha \quad \text{fr } \alpha, \beta \in A \text{ und } \alpha < \beta$$

so heit E der *direkte (induktive) Limes* von

$$(E_\alpha; r_{\alpha, \beta});$$

$$E := \lim_{\rightarrow} E_\alpha$$

2.2.1 Definition:

Seien $\vec{X}_1, \dots, \vec{X}_n$ unabhängige und identisch verteilte Zufallsvektoren, $n \in \mathbb{N}$, weiter seien $(P(\vec{X}^n) : n \in \mathbb{N})$ die daraus gebildeten Mengensysteme. Für alle $n \in \mathbb{N}$ läßt sich die folgende Abbildung definieren:

$$\left. \begin{aligned} r_{n,n+1}(B_{\underline{k}_i, k_i}(\vec{X}^n)) &:= B_{\underline{k}_i, k_i}(\vec{X}^{n+1}) \\ r_{n,n+1}(B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n)) &:= B_{\underline{k}_i, \bar{k}_i}(\vec{X}^{n+1}) \end{aligned} \right\} \quad i = 1, \dots, n$$

$$r_{n,n+1}(\mathbb{R}_p) := \mathbb{R}_p$$

$$r_{n,n+1}(\emptyset) := \emptyset ;$$

weiter sei für die Mengen $A, B \in P(\vec{X}^n)$ definiert:

$$r_{n,n+1}(A \Delta B) := r_{n,n+1}(A) \Delta r_{n,n+1}(B).$$

Um nachzuweisen, daß diese Abbildungen wohl definiert sind, beachte man, daß für Elemente $A, B \in P(\vec{X}^n)$, $n \in \mathbb{N}$ entweder gilt

$$\begin{aligned} &A \subset B \text{ oder } B \subset A \text{ oder } A \cap B = \emptyset \Rightarrow \\ &A \Delta B = \begin{cases} B \setminus A & \text{falls } A \subset B \\ A \setminus B & \text{falls } B \subset A \\ \mathbb{R}_p & \text{falls } A \cap B = \emptyset \end{cases} \end{aligned}$$

Es ist zu zeigen:

es gelte o.E. $A \subset B \Rightarrow r_{n,n+1}(A) \subset r_{n,n+1}(B)$, $n \in \mathbb{N}$

und $\emptyset = A \cap B \Rightarrow r_{n,n+1}(A) \cap r_{n,n+1}(B) = \emptyset$

$$\begin{aligned}
 \text{i)} \quad & \text{Sei o.E. } A := B_{\underline{k}_i, k_i}(\vec{X}^n) \text{ und} \\
 & B := B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n) \quad \implies \\
 & r_{n,n+1}(A) = B_{\underline{k}_i, k_i}(\vec{X}^{n+1}) \subset \\
 & B_{\underline{k}_i, \bar{k}_i}(\vec{X}^{n+1}) = r_{n,n+1}(B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n))
 \end{aligned}$$

$$\begin{aligned}
 \text{ii)} \quad & \text{Sei o.E. } A := B_{\underline{k}_i, k_j}(\vec{X}^n) \text{ und} \\
 & B := B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n) \quad \implies \\
 & r_{n,n+1}(B_{\underline{k}_i, k_i}(\vec{X}^n)) \cap r_{n,n+1}(B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n)) = \\
 & B_{\underline{k}_i, k_i}(\vec{X}^{n+1}) \cap B_{\underline{k}_i, \bar{k}_i}(\vec{X}^{n+1}) = \emptyset
 \end{aligned}$$

iii) Insgesamt gilt für

$$A = B_{\underline{k}_i, k_i}(\vec{X}^n) \text{ und}$$

$$B = B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n)$$

$$r_{n,n+1}(A \Delta B) = r_{n,n+1}(A) \Delta r_{n,n+1}(B)$$

$$\text{und für } A = B_{\underline{k}_i, k_i}(\vec{X}^n) \text{ und für } B = B_{\underline{k}_i, \bar{k}_i}(\vec{X}^n)$$

ebenfalls

$$r_{n,n+1}(A \Delta B) = r_{n,n+1}(A) \Delta r_{n,n+1}(B).$$

2.2.2 Korollar:

i) Für alle $n \in \mathbb{N}$ ist $(A(P(\vec{X}^n)); \Delta)$ eine Gruppe;
 $r_{n,n+1}$ ist somit ein Homomorphismus.

ii) Für alle $m, n \in \mathbb{N}$, $m > n$ ist

$$r_{n,m} := r_{m-1,m} \circ r_{m,m+1} \circ \dots \circ r_{n,n+1}$$

ein Homomorphismus.

iii) Aus i) und ii) folgt:

$(A(P(\vec{X}^n)); r_{n,m}; n, m \in \mathbb{N}; n < m)$ ist ein induktives System für μ -fast jede Realisation von $\vec{X}_1, \dots, \vec{X}_n$, wenn zudem $r_{n,n} := \text{id}$ gilt.

iv) Für jede absteigende Folge von Partitionen $(P_n: n \in \mathbb{N})$ ist $A((P_n): n \in \mathbb{N})$ versehen mit den Inklusionen

$$i_{n,m}: A(P_n) \longrightarrow A(P_m) \quad n < m, n, m \in \mathbb{N}$$

$$A \longrightarrow A$$

ein induktives System und $A((P_n: n \in \mathbb{N}))$ dessen induktiver Limes mit den kanonischen Abbildungen

$$i_n: A(P_n) \longrightarrow A((P_n: n \in \mathbb{N})) \quad n \in \mathbb{N}$$

$$A \longrightarrow A$$

Beweis:

Es genügt, iv) nachzuweisen; zu zeigen ist:

$$A((P_n: n \in \mathbb{N})) = \bigcup_{n \in \mathbb{N}} i_n(A(P_n)).$$

Es ist sofort einzusehen, daß

$$\bigcup_{n \in \mathbb{N}} i_n(A(P_n)) \subset A((P_n: n \in \mathbb{N})) \text{ gilt.}$$

Zum Beweis der "anderen Richtung" sei

$$A \in A((P_n: n \in \mathbb{N})), \text{ daraus folgt } A = \bigcup_{i=1}^1 A_i, A_i \in \mathcal{P}_{n_i}, n_i \in \mathbb{N}$$

$$n_* := \max_{i=1, \dots, l} n_i \implies A_i \in A(P_{n_*}), i=1, \dots, l$$

$$\bigcup_{i=1}^1 A_i \in A(P_{n_*}) \implies i_{n_*} \left(\bigcup_{i=1}^1 A_i \right) = \bigcup_{i=1}^1 i_{n_*} A_i \in \bigcup_{n \in \mathbb{N}} A(P_n) \quad //.$$

Nach diesen Vorbereitungen ist es jetzt möglich, schwache Konvergenz (im Sinne von stochastischer Konvergenz) von StEB's gegen einen entsprechenden Limes zu definieren.

2.2.3 Definition

$(\mathbb{R}_p; \mathcal{B}_p; \mu)$ sei der zugrunde liegende Wahrscheinlichkeitsraum, $p \geq 2$, μ sei stetig. Für μ -fast alle Realisationen $\vec{x}^n$ von $\vec{X}^n$, $n \in \mathbb{N}$, sei $(P(\vec{x}^n), r_{n,m}, n < m, n, m \in \mathbb{N})$ ein induktives System.

i) $(P(\vec{x}^n); r_{n,m}, n < m, n, m \in \mathbb{N})$ ist konvergent

: $\iff$

$(P(\vec{X}^n); r_{n,m}, n < m, n, m \in \mathbb{N})$ besitzt für

μ -fast alle Realisationen $\vec{x}^n$ von $\vec{X}^n$, $n \in \mathbb{N}$

einen induktiven Limes $\lim_{\leftarrow} P(\vec{x}^n)$

ii) Sei $(P_n: n \in \mathbb{N})$ eine Folge von Partitionen.

$(P(\vec{X}^n), r_{n,m}, n < m, n, m \in \mathbb{N})$ konvergiert

schwach gegen $A(P_n: n \in \mathbb{N})$: $\iff$

$(P(\vec{X}^n), r_{n,m}, n < m, n, m \in \mathbb{N})$ ist konvergent

und $\forall \epsilon > 0 \quad \forall n \in \mathbb{N} \quad \exists C_n \in \mathcal{C}_n(\vec{x}^n) \quad \exists A_n \in \mathcal{P}_n \quad \lim_{m \rightarrow \infty} W(\mu(r_{n,m}(C_n)) \Delta A_n) > \epsilon = 0$

2.2.4 Bemerkung

Aus Definition 2.2.3 und 2.2.2 folgt: Ist $(P_n: n \in \mathbb{N})$ eine absteigende Folge von Partitionen, dann konvergiert $(P(\vec{X}^n), r_{n,m}, n < m, n, m \in \mathbb{N})$ schwach gegen $A(P_n: n \in \mathbb{N})$, wenn für μ -fast alle induktiven Limes $\lim_{\leftarrow} P(\vec{x}^n)$ gilt.

$$\forall C \in \lim_{\leftarrow} P(\vec{x}^n) \quad \exists A \in \lim_{\leftarrow} A(P_n) \quad \mu(C \Delta A) = 0$$

Eine Folge von StEB's, die der *Definition 2.2.3* genügt, repräsentiert mit wachsendem Stichprobenumfang eine vorgegebene, für das anstehende Problem günstige Folge von Partitionen immer besser. Auch sind hierdurch, zumindest was die Bezeichnungsweisen anbelangt, die beiden Begriffe, die T. W. Anderson in seinen Konsistenzbetrachtungen nebeneinanderstellt - die Folge von Partitionen und die StEB's - in Beziehung gebracht. Was mit diesen Hilfsmitteln nicht möglich ist, ist nachzuweisen, daß T. W. Anderson's Folge von StEB's tatsächlich gegen die Partitionenfolge konvergiert, oder allgemeiner, die Existenz von konvergenten Partitionenfolgen zu zeigen. Um dazu in der Lage zu sein, ist eine andere Betrachtungsweise der StEB's angezeigt. Sie wird, wie im nachfolgenden Paragraphen entwickelt werden wird, darauf beruhen, eine Abbildung zu definieren, die den Elementen, die die einzelnen Teilmengen der StEB's bestimmen, die Wahrscheinlichkeit bezüglich eines stetigen Wahrscheinlichkeitsmaßes μ zuordnet.

§ 3

Definition der Statistically Equivalent Blocks durch Abbildungen

Wie im einleitenden Paragraphen bereits angekündigt, steht im Vordergrund dieses Abschnitts eine neue Definition der StEB's, die sich hauptsächlich Abbildungen bedient.

3.1 Vorbereitende Definitionen

Ausgegangen wird, wie immer in dieser Arbeit, von dem Verfahren StEB's zu bilden, wie es in § 1 vorgestellt wurde. Im einzelnen seien:

$\vec{X}_1, \dots, \vec{X}_n$ n Realisationen der stetigen, unabhängigen und identisch verteilten Zufallsvektoren $\vec{X}_1, \dots, \vec{X}_n$,

$(\mathbb{R}_p, \mathcal{B}_p, \mu)$, $p \geq 2$ der zugrundeliegende Wahrscheinlichkeitsraum, F_μ die Verteilungsfunktion von μ und f_μ , die nach Voraussetzung existierende Dichte. Weiter seien

$h_1, \dots, h_n$ n reellwertige, nicht notwendig verschiedene Funktionen, die bewirken, daß

$h_i(\vec{X}_j)$, $i=1, \dots, n$, $j=1, \dots, n$ stetige, reelle Zufallsvariable sind.

Seien $k := |\{h_1, \dots, h_n\}|$ und o. E. $h_1, \dots, h_k$ die entsprechenden Abbildungen, die zu einer meßbaren Funktion $H : \mathbb{R}_p \longrightarrow \mathbb{R}_k$,

$$\vec{x} \longrightarrow H(\vec{x}) = (h_1(\vec{x}), \dots, h_k(\vec{x}))$$

zusammengefaßt werden;

$$\pi : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$$

$$\pi(i) = k_i \quad i=1, \dots, n$$

sei die dem Verfahren zugrundeliegende Permutation.

Im weiteren Verlauf der Arbeit werden einige Größen und Mengen angesprochen werden, die bei der Konstruktion der StEB's, wie sie in § 1 vorgeführt wurde, verwendet werden; sie seien anschließend zur Erinnerung noch einmal kurz aufgeführt:

$$\begin{aligned} - \underline{k}_1 &:= 1, \quad \bar{k}_1 := n + 1 \\ \underline{k}_i &:= \max\{j \in \{0, k_1, \dots, k_{i-1}\} : j < k_i\} \\ \bar{k}_i &:= \min\{j \in \{k_1, \dots, k_{i-1}, n+1\} : j > k_i\} \quad i=2, \dots, n \end{aligned}$$

- aus der Stichprobe $\vec{X}_1, \dots, \vec{X}_n$ werden sukzessive mittels der Abbildungen $h_1, \dots, h_n$ die reellen Zufallsvariablen $Z_{k_1}, \dots, Z_{k_n}$ ermittelt; ihre Realisationen seien mit $z_{k_1}, \dots, z_{k_n}$ bezeichnet.

- darauf aufbauend wurden die Mengen

$$B_{0, k_1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) \leq z_{k_1}\}$$

$$B_{k_1, n+1}(\vec{X}) := \{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) > z_{k_1}\}$$

$$\left. \begin{aligned} B_{\underline{k}_i, k_i}(\vec{X}) &:= B_{\underline{k}_i, \bar{k}_i}(\vec{X}) \cap \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq z_{k_i}\} \\ B_{k_i, \bar{k}_i}(\vec{X}) &:= B_{\underline{k}_i, \bar{k}_i}(\vec{X}) \cap \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) > z_{k_i}\} \end{aligned} \right\} i=2, \dots, n$$

definiert; weiter gelte

$$B_{0, n+1}(\vec{X}) = \mathbb{R}_p$$

Im n-ten Schritt sind dann $n + 1$ disjunkte Mengen entstanden und es wurde für $i = 1, \dots, n+1$

$$B_{i-1, i}(\vec{X}) := B_i(\vec{X}) \text{ gesetzt.}$$

- Durch Integration der StEB's mit dem Wahrscheinlichkeitsmaß μ entstehen die Zufallsvariablen

$$V_{k_1} := \mu(B_{O, k_1}(\vec{X}))$$

$$U_{k_1} := V_{k_1}$$

$$V_{k_i} := \mu(B_{k_1, k_i}(\vec{X}))$$

$$U_{k_i} := \sum_{\substack{k_j \leq k_i \\ j \leq i}} V_{k_j}, \quad i = 1, \dots, n;$$

v_{k_i} und u_{k_i} bezeichnen die entsprechenden Realisationen.

Die sich jetzt anschließenden Betrachtungen gründen sich nicht mehr auf die "ganze" zugrunde liegende Stichprobe, sondern nur noch auf die durch das Verfahren ermittelten Zufallsgrößen $Z_{k_1}, \dots, Z_{k_n}$. Wie in § 1 dargestellt wurde, werden diese Zufallsvariablen aus $\vec{X}_1, \dots, \vec{X}_n$ mittels der Schnittfunktionen $h_1, \dots, h_n$, der Permutation π und der Zuordnung der Indices an die Schnittfunktionen (die angibt, in welchem Schritt welche Funktion verwendet wird) erstellt. Die Schnittfunktionen, die Permutation und die Reihenfolge der Verwendung der Schnittfunktionen sind jetzt als fest vorgegeben vorausgesetzt und werden selbst nicht mehr diskutiert.

Zunächst geht man dazu über, nicht mehr Zufallsvektoren sondern n-Tupel von p-dimensionalen Vektoren, auch mit $(\vec{x}_1, \dots, \vec{x}_n)$ bezeichnet und n-Tupel reeller Zahlen, wieder mit $(z_{k_1}, \dots, z_{k_n})$ bezeichnet, zu betrachten. Diese entstehen aus $\vec{x}_1, \dots, \vec{x}_n$ in Übereinstimmung mit dem Verfahren zur Bildung von StEB's mittels der Schnittfunktionen, der Permutation und der Zuordnung der Indices an die Schnittfunktionen.

Um das Verfahren, StEB's zu bilden, mit diesen festen Größen eindeutig durchführen zu können, muß man die Menge der in Betracht gezogenen $\vec{x}_1, \dots, \vec{x}_n$ einschränken, beispielsweise muß gelten $h_{k_1}(\vec{x}_i) \neq h_{k_1}(\vec{x}_j)$ für $i \neq j$; dies wird in der anschließenden Definition formalisiert werden.

Eine weitere Einschränkung, die für den Rest der Arbeit verbindlich ist, ist die, nur noch die p Projektionen auf die Koordinatenachsen als Schnittfunktionen zuzulassen, und zwar sollen, aus Gründen der Unterscheidbarkeit zweier Wahrscheinlichkeitsmaße, wie später ausgeführt werden wird, alle p Projektionen Verwendung finden. Der einfacheren Bezeichnungsweise wegen werden die Schnittfunktionen auch in diesem Fall mit $h_{k_1}, \dots, h_{k_n}$ angesprochen. Auf diese Einschränkungen könnte an manchen Stellen verzichtet werden, sie wurden trotzdem getroffen, da durch diese Maßnahme die Darstellungs- und Sichtweise der weiteren Ausführungen wesentlich erleichtert wird. Mit diesen Annahmen können jetzt einige Definitionen formuliert werden.

3.1.1 Definition

$$i) \quad \underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)}_{n\text{-mal}}_0 := \{ (\vec{x}_1, \dots, \vec{x}_n) \in \underbrace{\mathbb{R}_p \times \dots \times \mathbb{R}_p}_{n\text{-mal}} : \\ \prod_{r=1}^n (pr_i(\vec{x}_1) - pr_i(\vec{x}_r)) \neq 0, i = 1, \dots, p \}$$

- ii) die Abbildung, die jedem n -Tupel $(\vec{x}_1, \dots, \vec{x}_n) \in (\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0$ gemäß dem Verfahren StEB's zu bilden, ein n -Tupel $(z_{k_1}, \dots, z_{k_n}) \in \mathbb{R}_n$ zuordnet, sei mit S_π bezeichnet;
 $S_\pi((\vec{x}_1, \dots, \vec{x}_n)) = (z_{k_1}, \dots, z_{k_n})$.

Mit diesen Definitionen ist es jetzt möglich, eine Abbildung von $S_\pi((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0)$ in $[0;1]^n$ zu definieren

3.1.2 Definition

Sei μ ein stetiges Wahrscheinlichkeitsmaß auf $(\mathbb{R}_p; B_p)$, dann sei

$$F: S_\pi((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0) \longrightarrow [0;1]^n$$

$$F((z_{k_1}, \dots, z_{k_n})) := (F_1(z_{k_1}), F_2((z_{k_1}, z_{k_2})), \dots, F_n((z_{k_1}, \dots, z_{k_n})))$$

$$F_1(z_{k_1}) := \mu(B_{0, k_1}(\vec{x})) = u_{k_1}$$

$$F_i((z_{k_1}, \dots, z_{k_i})) := \sum_{\substack{k_j \leq k_i \\ j \leq i}} \mu(B_{k_j, k_j}(\vec{x})) = \sum_{\substack{k_j \leq k_i \\ j \leq i}} v_{k_j} = u_{k_i}$$

3.1.3 Bemerkung

- i) Die Abbildung S_π subsumiert die Schritte bei der Bildung von StEB's, die *nicht* vom Wahrscheinlichkeitsmaß μ , mit dem integriert wird, abhängen.
- ii) Die Abbildung F hängt von dem Wahrscheinlichkeitsmaß μ ab ¹⁾

1) Aus Gründen der einfacheren Schreibweise wurde darauf in der Bezeichnung nicht eingegangen. Sei z.B. $v \neq \mu$ ein weiteres Wahrscheinlichkeitsmaß, so wird dessen entsprechende Abbildung mit G bezeichnet werden.

3.2 Definition einer Umkehrabbildung in Anlehnung an den eindimensionalen Fall

Im eindimensionalen Fall wird als Umkehrabbildung einer Verteilungsfunktion F eines stetigen Wahrscheinlichkeitsmaßes folgendes definiert:

$$F^{-1}: (0;1) \longrightarrow \mathbb{R}$$

$$u \longrightarrow \inf\{x \in \mathbb{R} : F(x) = u\};$$

ist F streng monoton, so stimmt diese Definition mit der einer "gewöhnlichen" Umkehrabbildung überein. Umkehrabbildungen werden vor allem, um Unterschiede zwischen zwei Verteilungen μ und ν zu untersuchen, betrachtet. Zu diesem Zweck wird auch jetzt, im Falle höherer Dimensionen eine Umkehrabbildung definiert; ausgegangen wird hierbei von einer schon festgelegten Permutation π und vorgegebenen Schnittfunktionen $h_1, \dots, h_n$. Den weiteren Überlegungen wird nur noch eine geeignete (nicht notwendig echte) Teilmenge des $\mathbb{R}_n$, das Ziel der Abbildung S_π , zugrunde liegen.

Darüberhinaus wird die Stichprobe in ihrer ursprünglichen Form nicht mehr benötigt.

Zunächst einige Festlegungen:

- $D_0^a := \{(u_{(1)}, \dots, u_{(n)}) \in \mathbb{R}_n : 0 < u_{(1)} \leq \dots \leq u_{(n)} < 1\}$
- in Abhängigkeit der Permutation μ läßt sich die folgende Menge definieren:

$$D_\pi^a := \{(u_{k_1}, \dots, u_{k_n}) \in \mathbb{R}_n : u_{k_i} = \text{pr}_{k_i}((u_{(1)}, \dots, u_{(n)})), \\ i = 1, \dots, n, (u_{(1)}, \dots, u_{(n)}) \in D_0^a\}$$

D_π^a geht aus D_0^a durch "Umordnen" hervor; diese Umordnung ist, da das Gleichheitszeichen zugelassen ist, zunächst nicht eindeutig.

3.2.1 Definition:

Seien die Menge $S_{\pi}(\underbrace{(\mathbb{R}_p \dots \mathbb{R}_p)_o}_{n\text{-mal}}) \subset \mathbb{R}_n$ und die "Indices" $\underline{k}_i, \bar{k}_i, i = 1, \dots, n$, wie im vorangegangenen Abschnitt erklärt, $h_i \in \{pr_1, \dots, pr_p\}, i = 1, \dots, n$ und μ ein stetiges Wahrscheinlichkeitsmaß auf $(\mathbb{R}_p, B_p)$; dann sei

$$F^{-1}: D_{\pi}^a \longrightarrow S_{\pi}((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_o)$$

$$F^{-1}((u_{k_1}, \dots, u_{k_n})) := (F_1^{-1}(u_{k_1}), \dots, F_i^{-1}(u_{k_1}, \dots, u_{k_i}), \dots, F_n^{-1}(u_{k_1}, \dots, u_{k_n}))$$

$$F_1^{-1}(u_{k_1}) := \inf\{y \in \mathbb{R} : \mu(\{\vec{x} \in \mathbb{R}_p : h_{k_1}(\vec{x}) \leq y\}) = u_{k_1}\} =: z_{k_1}$$

$$F_i^{-1}((u_{k_1}, \dots, u_{k_i})) := \inf\{y \in \mathbb{R} : \mu(B_{\underline{k}_i, \bar{k}_i}(\vec{x}) \cap$$

$$\{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq y\}) = u_{k_i} - u_{\underline{k}_i}\} =: z_{k_i}$$

$$B_{\underline{k}_i, k_i}(\vec{x}) := \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq z_{k_i}\} \cap B_{\underline{k}_i, \bar{k}_i}(\vec{x})$$

Bei diesen Festlegungen und Definitionen wirken zwei Dinge störend: die zugelassenen Gleichheitszeichen in der Menge D_{π}^a und die Infimumbildung bei der Definition der Abbildung F^{-1} . Beides kann durch Einschränkungen an die Quelle $S_{\pi}((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_o)$ der Abbildung F , die durch das spezielle Wahrscheinlichkeitsmaß μ , das der Abbildung F zugrunde liegt, bestimmt ist, behoben werden.

1) Die Bezeichnungen sind, um die Analogie zu verdeutlichen, absichtlich wie in § 1 gewählt.

Man betrachte nicht die Menge $(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0$, sondern eine Menge, die auf dem kartesischen Produkt des Trägers Tr_μ der Dichte f_μ des Wahrscheinlichkeitsmaßes μ , aufbaut:

$$(\underbrace{\text{Tr}_\mu \times \dots \times \text{Tr}_\mu}_{n\text{-mal}})_0 := \{(\vec{x}_1, \dots, \vec{x}_n) \in \underbrace{\text{Tr}_\mu \times \dots \times \text{Tr}_\mu}_{n\text{-mal}} : \dots\}$$

$$\prod_{\substack{1, j \\ 1 \neq j}} (\text{pr}_i(\vec{x}_1) - \text{pr}_i(\vec{x}_j)) \neq 0, \quad i = 1, \dots, p\}$$

Weiter gilt:

$$F(S_\pi(\text{Tr}_\mu \times \dots \times \text{Tr}_\mu)_0) = D_\pi := \{(u_{k_1}, \dots, u_{k_n}) \in D_\pi^3 :$$

$$u_{k_i} \neq u_{k_j}, \quad i \neq j\}$$

Die Infimumbildung bei der Definition von F^{-1} wird überflüssig, wenn F eine injektive Abbildung ist. Die obige Einschränkung auf das kartesische Produkt der Trägermengen von f_μ reicht aus, um zu zeigen, daß F eine injektive Abbildung ist, wie das folgende Lemma zeigt.

3.2.2 Lemma:

Sei μ ein stetiges Wahrscheinlichkeitsmaß auf $(\mathbb{R}_p; \mathcal{B}_p)$, f_μ dessen Dichte, Tr_μ deren Trägermenge und F die zu μ assoziierte Abbildung und für die Schnittfunktionen gelte

$$h_i \in \{\text{pr}_1, \dots, \text{pr}_p\}, \quad i = 1, \dots, n,$$

dann ist F eine bijektive Abbildung.

Beweis:

Zunächst einige Schreibweisen, die den Beweis übersichtlicher gestalten sollen:

$$\vec{z} := (z_{k_1}, \dots, z_{k_n}), \quad \vec{y} := (y_{k_1}, \dots, y_{k_n}),$$

$$\vec{y}, \vec{z} \in S_n((\text{Tr}_\mu \times \dots \times \text{Tr}_\mu) \circ)$$

$$B_{\underline{k}_i, \bar{k}_i}(\vec{z}) \text{ und } B_{\underline{k}_i, \bar{k}_i}(\vec{y}), \quad i = 1, \dots, n$$

seien die hieraus gebildeten Mengen.

Es sei

$$F(\vec{y}) = F(\vec{z})$$

$$\text{Annahme: } \vec{z} \neq \vec{y}$$

Sei hierzu k_i der erste Index, für den gilt

$$z_{k_i} \neq y_{k_i}, \quad \text{o.E. } z_{k_i} < y_{k_i};$$

$$\begin{aligned} B_{\underline{k}_i, k_i}(\vec{y}) &= B_{\underline{k}_i, \bar{k}_i}(\vec{y}) \cap (\{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq z_{k_i}\} \cup \\ &\quad \{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq y_{k_i}\}) \\ &= B_{\underline{k}_i, \bar{k}_i}(\vec{y}) \cap (\{\vec{x} \in \mathbb{R}_p : h_{k_i}(\vec{x}) \leq z_{k_i}\} \cup \\ &\quad \{\vec{x} \in \mathbb{R}_p : z_{k_i} < h_{k_i}(\vec{x}) \leq y_{k_i}\}) \\ &= B_{\underline{k}_i, k_i}(\vec{z}) \cup (B_{\underline{k}_i, \bar{k}_i}(\vec{y}) \cap \{\vec{x} \in \mathbb{R}_p : z_{k_i} < h_{k_i}(\vec{x}) \leq y_{k_i}\}). \end{aligned}$$

Da $\mu(B_{\underline{k}_i, k_i}(\vec{y})) = \mu(B_{\underline{k}_i, k_i}(\vec{z}))$ gilt, folgt

$$\mu(B_{\underline{k}_i, \bar{k}_i}(\vec{y}) \cap \{\vec{x} \in \mathbb{R}_p : z_{k_i} < h_{k_i}(\vec{x}) \leq y_{k_i}\}) = 0.$$

Weiter sei o.E.

$$C := B_{\underline{k}_i, \bar{k}_i}(\vec{y}) \cap \{\vec{x} \in \mathbb{R}_p : x_{k_i} < h_{k_i}(\vec{x}) \leq y_{k_i}\};$$

C ist ein Quader in $\mathbb{R}_p$; o.E. gelte für $r \leq p-1$:

$$C = (-\infty; \infty) \times \dots \times (-\infty; \infty) \times (x_{k_i}; y_{k_i}] \times (a_1; b_1] \times \dots \times (a_r; b_r]$$

$$\Rightarrow \mu(C) = \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} \int_{x_{k_i}}^{y_{k_i}} \int_{a_1}^{b_1} \dots \int_{a_r}^{b_r} f((y_1, \dots, y_r, \dots, y_p)) dy_1 \dots dy_r \dots dy_p$$

$$= \int_{x_{k_i}}^{y_{k_i}} \int_{a_1}^{b_1} \dots \int_{a_r}^{b_r} f_R((y_1, \dots, y_{r+1})) dy_1 \dots dy_r dy_{r+1} = 0$$

$$\Rightarrow \int_{a_1}^{b_1} \dots \int_{a_r}^{b_r} f_R((y_1, \dots, y_{r+1})) dy_1 \dots dy_r = 0$$

$$\Rightarrow \int_{-\infty}^{+\infty} \int_{a_1}^{b_1} \dots \int_{a_r}^{b_r} f_R((y_1, \dots, y_{r+1})) dy_1 \dots dy_r dy_{r+1}$$

$$= \int_{a_1}^{b_1} \dots \int_{a_1}^{b_r} f_R((y_1, \dots, y_r)) dy_1 \dots dy_r = 0$$

$$\Rightarrow \mu(B_{\underline{k}_i, \bar{k}_i}(\vec{y})) = \mu(B_{\underline{k}_i, \bar{k}_i}(z)) = 0 \Rightarrow$$

$$u_{\underline{k}_i}^- - u_{\underline{k}_i} = 0 \Rightarrow u_{\underline{k}_i}^- = u_{\underline{k}_i};$$

dies ist jedoch ein Widerspruch zur Annahme. //

Die Voraussetzung an die Schnittfunktionen, Projektionen zu sein, kann dahin erweitert werden, daß gefordert wird, die verschiedenen unter diesen Abbildungen können zu einer Bijektion zusammengefaßt werden. Da dies aber für die weiteren Betrachtungen ohne Belang ist, wird auf eine Ausführung dieser Verallgemeinerung verzichtet.

3.3 Differenzierbarkeit der Abbildungen F und F^{-1}

Dieser Abschnitt ist, wie sich im Folgenden erweisen wird, vornehmlich von technischem Interesse. Von nun an sind, wenn nicht ausdrücklich etwas anderes definiert wird, die nachstehenden Voraussetzungen getroffen:

- die Schnittfunktionen sind die Projektionen des $\mathbb{R}_p$ auf $\mathbb{R}$,
- das zugrundeliegende Wahrscheinlichkeitsmaß μ auf $(\mathbb{R}_p, \mathcal{B}_p)$ ist stetig mit einer auf $\mathbb{R}_p$ positiven und stetigen Dichte f_μ und der Verteilungsfunktion F_μ .

Betrachtet man nun die Abbildung

$$F: S_\pi \left(\underbrace{(\text{Tr}_\mu \times \dots \times \text{Tr}_\mu)_0}_{n\text{-mal}} \right) \longrightarrow D_\pi,$$

so ist sie, wie im vorangegangenen Abschnitt gezeigt wurde, bijektiv. Um die Differenzierbarkeit der Abbildungen F und F^{-1} zu zeigen, nützt man deren Abhängigkeit von der Verteilungsfunktion F_μ aus; dieser Zusammenhang sei im Folgenden kurz erläutert:

Es ist

$$F_i((z_{k_1}, \dots, z_{k_i})) = \sum_{j \geq i} \mu(B_{k_j, k_j}(\vec{x})).$$

$B_{k_j, k_j}(\vec{x})$ ist ein Quader in $\mathbb{R}_p$, dessen Eckpunkte nur die Koordinaten $-\infty, z_{k_1}, \dots, z_{k_j}$ haben können; $\mu(B_{k_j, k_j}(\vec{x}))$ ist durch die Verteilungsfunktion F_μ darstellbar.

Betrachtet man die partiellen Ableitungen

$$\frac{\partial}{\partial z_{k_j}} F_i((z_{k_1}, \dots, z_{k_i})),$$

so sind diese nur für $j \leq i$ ungleich Null, es gilt

$$\frac{\partial}{\partial z_{k_j}} F_i((z_{k_1}, \dots, z_{k_i})) = \sum_{k_j \leq k_1 \leq k_i} \frac{\partial}{\partial z_{k_j}} \mu(B_{k_1, k_1}(\vec{x}));$$

hiervon sind nur die Summanden ungleich Null, in denen $B_{k_1, k_1}(\vec{x})$ Eckpunkte mit einer Koordinate z_{k_j} besitzt.

Schreibt man hiervon wieder $\mu(B_{k_1, k_1}(\vec{x}))$ mittels F_μ und sei o.E.

$$B_{k_1, k_1}(\vec{x}) := \{\vec{z} \in \mathbb{R}_p :$$

$$z_i^u < z_i \leq z_i^o, z_i^u \in \mathbb{R} \cup \{-\infty\}, z_i^o \in \mathbb{R}, i = 1, \dots, p\},$$

so ergibt dies

$$\mu(B_{k_1, k_1}(\vec{x})) = F_\mu((z_1^o, \dots, z_p^o)) -$$

$$\sum_{i=1}^n F_\mu((z_1^o, \dots, z_{i-1}^o, z_1^u, z_{i+1}^o, \dots, z_p^o)) +$$

$$\sum_{j=2}^n \sum_{i=1}^j F_\mu((z_1^o, \dots, z_i^u, \dots, z_j^u, \dots, z_p^o)) - \dots$$

$$+ (-1)^n F_\mu((z_1^u, \dots, z_p^u)).$$

Diese Darstellung zeigt, daß es genügt, die Differenzierbarkeit von F_μ in Punkten $\vec{z} \in \mathbb{R}_p$ nachzuweisen, deren Komponenten der Menge

$$\{-\infty, z_{k_1}, \dots, z_{k_i}\}$$

entstammen.

Zum Nachweis der Stetigkeit der partiellen Ableitungen kann man sich, in analog begründeter Weise, auch auf die Verteilungsfunktion F_μ zurückziehen.

3.3.1 Satz:

Die Abbildung $F((z_{k_1}, \dots, z_{k_n}))$ ist λ_p -f.ü. stetig differenzierbar.

Beweis:

Zu zeigen ist:

- die partiellen Ableitungen

$$f^{i,j}((z_{k_1}, \dots, z_{k_n})) := \frac{\partial_i F((z_{k_1}, \dots, z_{k_n}))}{\partial z_{k_j}}$$

existieren für $i=1, \dots, n$ und $j=1, \dots, n$ λ_n -f.ü.;

- $f^{i,j}((z_{k_1}, \dots, z_{k_n}))$ ist für $i=1, \dots, n$ und $j=1, \dots, n$ λ_n -f.ü. stetig.

$$F_j((z_{k_1}, \dots, z_{k_j})) = u_{k_j} = \\ u_{k_j} + \mu(\{\vec{x} \in \mathbb{R}_p : h_{k_j}(\vec{x}) \leq z_{k_j}\} \cap B_{\vec{x}_j, \vec{k}_j}(\vec{x}))$$

somit ist nur die Differenzierbarkeit des zweiten Summanden nachzuweisen; dies reduziert sich, wie erläutert, zum Nachweis der Differenzierbarkeit der Verteilungsfunktion F_μ an den Stellen $\vec{x} \in \overline{\mathbb{R}}_p$.

- i) Wegen des *Satzes von Fubini* genügt es, die Nachweise für die Komponente x_1 durchzuführen:

$$F_\mu(x_1, \dots, x_p) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_p} f_\mu(y_1, \dots, y_p) dy_p \dots dy_1$$

somit ist

$$\frac{\partial}{\partial x_1} F_\mu(x_1, \dots, x_p) = \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} f_\mu(x_1, y_2, \dots, y_p) dy_p \dots dy_2$$

folglich existieren die partiellen Ableitungen nach $x_1, \dots, x_p$.

- ii) Wie angekündigt, bleibt die Stetigkeit der partiellen Ableitungen zu zeigen; da auch die Summe von stetigen Abbildungen wieder stetig ist, genügt es, die Stetigkeit von

$$\frac{\partial}{\partial x_1} F_\mu(x_1, \dots, x_p), i=1, \dots, p \text{ zu beweisen.}$$

Hierzu sei, wie oben, o.E. $i = 1$.

$$\frac{\partial}{\partial x_1} F_\mu(x_1, \dots, x_p) =: T((x_1, x_2, \dots, x_p)).$$

Zunächst wird gezeigt:

Seien f_u und f_u^1 stetig in x_1 , dann gilt für alle¹⁾

Folgen $(x_n^1: n \in \mathbb{N})$ mit $\lim_{n \rightarrow \infty} x_n^1 = x_1$

$$\lim_{n \rightarrow \infty} T((x_n^1; x_2, \dots, x_p)) = T((x_1; x_2, \dots, x_p));$$

wobei dem Konvergenzbegriff die in $\mathbb{R}$ übliche Metrik zugrunde liege.

$$\text{iii) } t_n: \mathbb{R}_{p-1} \longrightarrow \mathbb{R}$$

$$t_n((y_2, \dots, y_p)) := f_u((x_n^1, y_2, \dots, y_p)), \quad n \in \mathbb{N}$$

$$t_o: \mathbb{R}_{p-1} \longrightarrow \mathbb{R}$$

$$t_o((y_2, \dots, y_p)) := f_u((x_1, y_2, \dots, y_p))$$

wegen der Stetigkeit von f_u in x_1 gilt:

$$\lim_{n \rightarrow \infty} t_n((y_2, \dots, y_p)) = \lim_{n \rightarrow \infty} f_u((x_n^1, y_2, \dots, y_p)) =$$

$$f_u((x_1, y_2, \dots, y_p)) = t_o((y_2, \dots, y_p)) \implies$$

t_n konvergiert punktweise gegen t_o ;

weiter gilt für λ_1 -f.a. $x_1 \in \mathbb{R}$, $\lim_{n \rightarrow \infty} x_n^1 = x_1$

$$\lim_{n \rightarrow \infty} E(t_n((y_2, \dots, y_p))) :=$$

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_u((x_n^1, y_2, \dots, y_p)) dy_2 \dots dy_p =$$

$$\lim_{n \rightarrow \infty} f_u^1(x_n^1) = f_u^1(x_1) =$$

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_u((x_1, y_2, \dots, y_p)) dy_p \dots dy_2 =:$$

$$E(t_o((y_2, \dots, y_p))).$$

1) f_u^1 bezeichnet die Randdichte

iv) Sei $\lim_{n \rightarrow \infty} x_n^1 = x_1$, so ist

$$T((x_n^1; x_2, \dots, x_p)) = \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} f_{\mu}((x_n^1, y_2, \dots, y_p)) dy_p \dots dy_2$$

—→

$$| T((x_n^1; x_2, \dots, x_p)) - T((x_1; x_2, \dots, x_p)) | =$$

$$\left| \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} (f_{\mu}((x_n^1, y_2, \dots, y_p)) - f_{\mu}((x_1, y_2, \dots, y_p))) dy_p \dots dy_2 \right| \leq$$

$$\int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} |f_{\mu}((x_n^1, y_2, \dots, y_p)) - f_{\mu}((x_1, y_2, \dots, y_p))| dy_p \dots dy_2 \leq$$

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} |f_{\mu}((x_n^1, y_2, \dots, y_p)) - f_{\mu}((x_1, y_2, \dots, y_p))| dy_p \dots dy_2 =$$

$$E(|t_n - t_0|) \quad (n \in \mathbb{N})$$

nach dem Lemma von Scheffé¹⁾ konvergiert $E(|t_n - t_0|)$

gegen null und somit folgt

$$\lim_{n \rightarrow \infty} T((x_n^1; x_2, \dots, x_p)) = T((x_1; x_2, \dots, x_p)) \quad \text{für } \lambda_1\text{-f.a. } x_1 \in \mathbb{R}.$$

Die Punkte iii) und iv) sind als Hilfsmittel zum Nachweis der Stetigkeit von

$$\frac{\partial}{\partial x_1} F_{\mu}((x_1, \dots, x_p)) \quad \text{zu verstehen. Um diesen Nach-}$$

weis zu führen sei

1) Gänssler, P., Stute, W.: Wahrscheinlichkeitstheorie, Springer-Verlag, Berlin, Heidelberg, New York, p. 37, (1977).

$\vec{x} := (x_1, \dots, x_p) \in \mathbb{R}_p,$

f_μ stetig in $\vec{x}$ und

$((x_n^1, \dots, x_n^p))$ eine Folge in $\mathbb{R}_p$, die bzgl. der gewöhnlichen Metrik auf $\mathbb{R}_p$ gegen $\vec{x}$ konvergiere.

Hiermit gilt dann:

$$\begin{aligned} & |T((x_1; x_2, \dots, x_p)) - T((x_n^1; x_n^2, \dots, x_n^p))| = \\ & |T((x_1; x_2, \dots, x_p)) - T((x_n^1; x_2, \dots, x_p)) + \\ & \quad T((x_n^1; x_2, \dots, x_p)) - T((x_n^1; x_n^2, \dots, x_n^p))| \leq \\ & |T((x_1; x_2, \dots, x_p)) - T((x_n^1; x_2, \dots, x_p))| + \\ & |T((x_n^1; x_2, \dots, x_p)) - T((x_n^1; x_n^2, \dots, x_n^p))|. \end{aligned}$$

Da die Stetigkeit von $T((x_1; x_2, \dots, x_p))$ in x_1 bereits gezeigt wurde, genügt es nachzuweisen, daß der zweite Summand

$$|T((x_n^1; x_2, \dots, x_p)) - T((x_n^1; x_n^2, \dots, x_n^p))|$$

gegen Null konvergiert.

Zunächst vier Definitionen:

$$c := \sup_{\vec{x} \in \mathbb{R}_p} f_\mu(\vec{x}),$$

$$D_{x_2 \Delta x_n}^{+2} := \left(\bigcap_{i=2}^p \text{pr}_i^{-1}((-\infty; x_i]) \right) \Delta \left(\bigcap_{i=2}^p \text{pr}_i^{-1}((-\infty; x_n^i]) \right)$$

$$u_n^i := \min\{x_1, x_n^i\}$$

$$i=2, \dots, n$$

$$o_n^i := \max\{x_1, x_n^i\}$$

Mit diesen Festlegungen gilt dann die folgende Abschätzung:

$$\begin{aligned} & \left| \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_p} f_\mu((x_n^1; y_2, \dots, y_p)) dy_p \dots dy_2 - \right. \\ & \left. \int_{-\infty}^{x_n^2} \dots \int_{-\infty}^{x_n^p} f_\mu((x_n^1; y_2, \dots, y_p)) dy_p \dots dy_2 \right| \leq \end{aligned}$$

$$\int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} f_{\mu}((x_n^1; y_2, \dots, y_p)) \cdot I_{D_{x_2 \Delta x_n}^+}(y_2, \dots, y_p) dy_p \dots dy_2 \\ \leq c \cdot \lambda_{p-1}(D_{x_2 \Delta x_n}^+) .$$

Da weiter

$$D_{x_2 \Delta x_n}^+ \subset \left(\bigcap_{i=2}^p \text{pr}_i^{-1}((-\infty; o_n^i]) \right) \setminus \left(\bigcap_{i=2}^p \text{pr}_i^{-1}((-\infty; u_n^i]) \right) =: D_n$$

und $D_n \downarrow \bigcap_{i=2}^p \text{pr}_i^{-1}(\{x_i\})$ gilt, folgt

$$\lim_{n \rightarrow \infty} c \cdot \lambda_{p-1}(D_{x_2 \Delta x_n}^+) = 0 ;$$

hiermit ist dann die Stetigkeit der partiellen Ableitungen bewiesen.

Zusammenfassend gilt:

Die Abbildungen $F_i((z_{k_1}, \dots, z_{k_i}))$ sind auf

$S_{\pi}((\text{Tr}_{\mu}^x \dots \times \text{Tr}_{\mu}^x)_0)$ stetig differenzierbar, somit ist F stetig differenzierbar und besitzt, bei geeigneter Einschränkung des Bildraumes, eine inverse Abbildung F^{-1} , die nach dem Satz von der inversen Abbildung stetig ist und eine Funktionaldeterminante, die in allen Punkten ungleich Null ist, besitzt. //

Im Fall endlicher Stichprobenumfänge wird die Umkehrfunktion benutzt werden, um Unterschiede zweier Wahrscheinlichkeitsverteilungen μ und ν zu charakterisieren; diese Aufgabe wird Gegenstand des folgenden Abschnitts sein.

3.4 Betrachtungen bei festem Stichprobenumfang n

Das Verfahren StEB's zu bilden unterliegt auch in diesem Abschnitt einigen einschränkenden Annahmen. Im Gegensatz zu den vorangegangenen Betrachtungen werden jetzt zwei Wahrscheinlichkeitsmaße μ und ν betrachtet. Ziel ist, Unterschiede zu beschreiben, die dadurch entstehen, daß die StEB's z.B. mit dem Wahrscheinlichkeitsmaß μ gebildet werden, jedoch ν zugrunde lag. Beide Maße induzieren Abbildungen: μ die Abbildung F und ν die Abbildung G

$$F: S_{\pi}(\underbrace{(\text{Tr}_{\mu} \times \dots \times \text{Tr}_{\mu})_0}_{n\text{-mal}}) \longrightarrow D_{\pi}$$

$$G: S_{\pi}(\underbrace{(\text{Tr}_{\nu} \times \dots \times \text{Tr}_{\nu})_0}_{n\text{-mal}}) \longrightarrow D_{\pi}$$

Nun sind die Trägermengen der Dichten f_{μ} und f_{ν} i.A. nicht identisch. so daß eine Abbildung

$$F \circ G^{-1}: D_{\pi} \longrightarrow D_{\pi}$$

eventuell nicht definiert ist, wenn z.B.

$$\text{Tr}_{\mu} \cap \text{Tr}_{\nu} = \emptyset \text{ ist.}$$

Diese Überlegung führt zu zwei einschränkenden Annahmen an die in Betracht gezogenen Wahrscheinlichkeitsmaße, die jedoch im wesentlichen der leichteren Darstellung dienen und teilweise mit geringfügigen Modifikationen an den betreffenden Aussagen fallengelassen werden können. Diese Einschränkungen sind:

- Die stetigen Wahrscheinlichkeitsmaße besitzen Dichten, die auf ganz $\mathbb{R}_p$ positiv und stetig sind. Ω bezeichne die Menge dieser Wahrscheinlichkeitsmaße.

Gegenstand dieses und des folgenden Abschnittes ist es, zu untersuchen, ob aus der Gleichheit der Verteilung von $F((Z_{k_1}, \dots, Z_{k_n}))$ und $G((Z_{k_1}, \dots, Z_{k_n}))$ auch die Gleichheit von μ und ν gefolgert werden kann. Ein erster Schritt hierzu ist die Untersuchung der Verteilung $F \circ G^{-1}((U_{k_1}, \dots, U_{k_n}))$, wobei die Dichte von $(U_{k_1}, \dots, U_{k_n})$ die Gestalt

$n! I_{D_\pi}((u_{k_1}, \dots, u_{k_n}))$ besitzt.

3.4.1 Lemma:

Seien $\mu, \nu \in \Omega$ und F und G deren assoziierte Abbildungen, dann sind die Abbildungen

$$F \circ G^{-1}: D_\pi \longrightarrow D_\pi \text{ und}$$

$$G \circ F^{-1}: D_\pi \longrightarrow D_\pi$$

Bijektionen,

Beweis:

$$F: S_\pi(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_{n\text{-mal}}}_p) \longrightarrow D_\pi$$

$$G: S_\pi(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_{n\text{-mal}}}_p) \longrightarrow D_\pi$$

Nach Lemma 3.2.2 sind F und G Bijektionen und folglich auch

$$F^{-1}: D_{\pi} \longrightarrow S_{\pi}(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_o}_{n\text{-mal}}) \text{ und}$$

$$G^{-1}: D_{\pi} \longrightarrow S_{\pi}(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_o}_{n\text{-mal}}) \quad //.$$

Die Abbildung $F \circ G^{-1}$ wird dazu benützt werden, die Verteilung des Zufallsvektors

$$(W_{k_1}, \dots, W_{k_n}) := F \circ S_{\pi}(\vec{X}_1, \dots, \vec{X}_n)$$

zu untersuchen, wenn die $\vec{X}_i$, $i = 1, \dots, n$ die Verteilung ν besitzen, die StEB's aber bzgl. μ integriert werden.

3.4.2 Lemma

Seien $\mu, \nu \in \Omega$, F und G deren assoziierte Abbildungen und $\vec{X}_1, \dots, \vec{X}_n$ unabhängige Zufallsvektoren, die die Verteilung ν besitzen; die StEB's seien mit μ integriert worden, dann besitzt die Verteilung von

$$(W_{k_1}, \dots, W_{k_n}) = F \circ S_{\pi}(\vec{X}_1, \dots, \vec{X}_n) = F((Z_{k_1}, \dots, Z_{k_n}))$$

die Dichte

$$g(\vec{w}) = \begin{cases} \left| \frac{\partial}{\partial \vec{w}} (F \circ G^{-1})^{-1}(\vec{w}) \right| n! & \text{falls } \vec{w} \in D_{\pi} \\ 0 & \text{sonst.} \end{cases}$$

Beweis:

Sei $\vec{u} := (u_{k_1}, \dots, u_{k_n}) = G((z_{k_1}, \dots, z_{k_n}))$ wobei

$$(z_{k_1}, \dots, z_{k_n}) = S_{\pi_n}((\vec{x}_1, \dots, \vec{x}_n));$$

$\vec{u}$ hat die Dichte

$$f(\vec{u}) = \begin{cases} n! & \text{falls } \vec{u} \in D_{\pi} \\ 0 & \text{sonst} \end{cases}$$

es gilt somit, $\vec{w} = \text{FoG}^{-1}(\vec{u})$, daraus folgt

$$g(\vec{w}) = \begin{cases} |\det \frac{\partial}{\partial \vec{w}} (\text{FoG}^{-1})^{-1}(\vec{w})| n! & \text{falls} \\ & (\text{FoG}^{-1})^{-1}(\vec{w}) \in D_{\pi} \\ 0 & \text{sonst} \end{cases}$$

da weiter gilt: $(\text{FoG}^{-1})^{-1}(\vec{w}) \in D_{\pi} \iff \vec{w} \in D_{\pi}$

folgt die Behauptung sofort. //

In Hinblick auf testtheoretische Fragestellungen besagt obiger Satz: $f((u_{k_1}, \dots, u_{k_n}))$ ist die Dichte von $(U_{k_1}, \dots, U_{k_n})$, wenn die Verteilung u , mit der "transformiert" wurde (Hypothese) falsch ist und v die "wahre" Verteilung ist. Folgendes Korollar ist sofort einzusehen:

3.4.3 - Korollar:

Unter den Voraussetzungen von 3.4.2 gilt:

$$\det \frac{\partial}{\partial \vec{u}} (\text{FoG}^{-1})^{-1}(\vec{u}) = 1 \iff f(\vec{u}) = g(\vec{u})$$

Folgender Satz beantwortet nun die Frage, ob aus

$$\left| \det \frac{\partial}{\partial \vec{w}} (F \circ G^{-1})^{-1}(\vec{w}) \right| = 1 \quad \text{auch}$$

$F = G$ gefolgert werden kann.

3.4.4 Satz:

Unter den Voraussetzungen von 3.4.2 an μ gilt:

$$\left| \det \frac{\partial}{\partial \vec{u}} (F \circ G^{-1})^{-1}(\vec{u}) \right| = 1 \iff F = G.$$

Beweis:

" $\Leftarrow$ ": Diese Richtung ist sofort einzusehen.

" $\Rightarrow$ ": Zur Wiederholung sei bemerkt:

$F \circ G^{-1}$ ist eine Bijektion von D_π in D_π ; die Dichte f der Verteilung des n -dimensionalen Zufallsvektors $\vec{U}$ habe die Gestalt

$$f(\vec{u}) = \begin{cases} n! & \text{falls } \vec{u} \in D_\pi \\ 0 & \text{sonst} \end{cases}$$

sei nun:

$$\vec{T} := F \circ G^{-1}, \quad \vec{T} = (T_1, \dots, T_n)$$

$$\vec{T}_i := (T_1, \dots, T_i), \quad i=1, \dots, n, \quad \vec{T}_i \text{ ist eine}$$

Bijektion von $(pr_1 \times \dots \times pr_i)(D_\pi)$ in $(pr_1 \times \dots \times pr_i)(D_\pi)$;

hierzu beachte man, daß T_j für $1 \leq j \leq i$ nur von

$u_{k_1}, \dots, u_{k_j}$ abhängt.

$$\vec{u}_i := (u_{k_1}, \dots, u_{k_i})$$

$$\vec{w} := (w_{k_1}, \dots, w_{k_n})$$

$$\vec{w}_i := (w_{k_1}, \dots, w_{k_i})$$

Weiter sei

$f_i((u_{k_1}, \dots, u_{k_i}))$ die Randdichte von $(u_{k_1}, \dots, u_{k_i})$,

$g_i((w_{k_1}, \dots, w_{k_i}))$ die Randdichte von $(w_{k_1}, \dots, w_{k_i})$,

$f_i^B(u_{k_i} | u_{k_1}, \dots, u_{k_{i-1}})$ und $g_i^B(w_{k_i} | w_{k_1}, \dots, w_{k_{i-1}})$

die entsprechenden bedingten Dichten.

$$i) \vec{u} \in \mathbb{R}_n, \vec{w} := \vec{T}(\vec{u}) \implies$$

$$g(\vec{w}) = \left| \det \frac{\partial \vec{T}^{-1}(\vec{w})}{\partial \vec{w}} \right| f(\vec{T}^{-1}(\vec{w}))$$

$$= n! \cdot I_D(\vec{T}^{-1}(\vec{w})) = n! \cdot I_D(\vec{u}) = f(\vec{u})$$

$$ii) \quad i) \implies f_1(w_{k_1}) = g_1(w_{k_1}) \quad \text{und} \\ g_1(w_{k_1}) = |T_1^{-1}(w_{k_1})| f_1(T_1^{-1}(w_{k_1})) \implies$$

$$g_1(w_{k_1}) = |T_1^{-1}(w_{k_1})| f_1(T_1^{-1}(w_{k_1})) \\ = |T_1^{-1}(w_{k_1})| g_1(T_1^{-1}(w_{k_1}))$$

$$\text{falls } w_{k_1} \in (0;1)$$

für $z \in (0;1)$ gilt

$$\int_0^z g_1(w_{k_1}) dw_{k_1} = \int_0^z |T_1^{-1}(w_{k_1})| g_1(T_1^{-1}(w_{k_1})) dw_{k_1} = \int_0^{T_1(z)} g_1(u_{k_1}) du_{k_1} \implies$$

$$F(z) = z, \quad z \in (0;1) \quad T_1 = \text{id.}$$

iii) Sei $2 \leq i \leq n$, $T_{i-1}(\vec{u}_{i-1}) = \vec{u}_{i-1}$ und

$$\begin{aligned} T_{u_{i-1}}^+ : (u_{k_i}^-; u_{k_i}^-] &\longrightarrow (\vec{u}_{k_i}^+; u_{k_i}^-] \\ u_{k_i} &\longrightarrow T_i((\vec{u}_{i-1}, u_{k_i})); \end{aligned}$$

$T_{u_{i-1}}^+$ ist bijektiv, so ist

$$\begin{aligned} g_i^B(w_{k_i} | \vec{w}_{i-1}) &= g_i^B(w_{k_i} | \vec{u}_{i-1}) = \\ \left| \frac{\partial}{\partial \vec{w}_{k_i}} T_{u_{i-1}}^{-1}(w_{k_i}) \right| \cdot f_i^B(T_{u_{i-1}}^{-1}(w_{k_i}) | \vec{u}_{i-1}) &= \\ \left| \frac{\partial}{\partial w_{k_i}} T_{u_{i-1}}^{-1}(w_{k_i}) \right| \cdot g_i^B(T_{u_{i-1}}^{-1}(w_{k_i}) | \vec{u}_{i-1}) &> 0. \end{aligned}$$

Sei weiter $z \in (u_{k_i}^-; u_{k_i}^-]$, so folgt

$$\begin{aligned} \int_{u_{k_i}^-}^z g_i^B(w_{k_i} | \vec{u}_{i-1}) dw_{k_i} &= \\ \int_{u_{k_i}^-}^z \left| \frac{\partial}{\partial w_{k_i}} T_{u_{i-1}}^{-1}(w_{k_i}) \right| g_i^B(T_{u_{i-1}}^{-1}(w_{k_i}) | \vec{u}_{i-1}) dw_{k_i} &= \end{aligned}$$

$$\begin{aligned} T_{u_{i-1}}^+(z) &= \\ \int_{u_{k_i}^-} g_i^B(u_{k_i} | \vec{u}_{i-1}) du_{k_i} &\implies \end{aligned}$$

$$T_{u_{i-1}}^+(z) = z \implies$$

$$T_i((u_{k_1}, \dots, u_{k_i})) = (u_{k_1}, \dots, u_{k_i}).$$

Aus i), ii), iii) folgt

$T = F \circ G^{-1}$ ist die identische Abbildung auf D_π
und somit gilt $F = G$.

//.

Für weitere Betrachtungen erweist sich die nachfolgend definierte Abbildung F_+ oft als günstiger. Seien zunächst

$$- t: D_\pi \longrightarrow \mathbb{R}_+^n$$

$$(v_1, \dots, v_n) := t((u_{k_1}, \dots, u_{k_n})) :=$$

$$(u_{k_1} - u_{k_1-1}, \dots, u_{k_n} - u_{k_n-1})$$

$$- D_+ := \{\vec{v} \in \mathbb{R}_+^n : 0 < v_i < 1, i=1, \dots, n, \sum_{i=1}^n v_i < 1\}$$

t ist eine in beiden Richtungen differenzierbare Bijektion, deren Funktionaldeterminante dem Betrage nach 1 ist; weiter ist

$$t(D_\pi) = D_+.$$

hiermit ist die Abbildung

$$F_+ := t \circ F: S_\pi(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p) \circ}_{n\text{-mal}}) \longrightarrow D_+$$

wohldefiniert.

3.4.5 Korollar:

Seien $\mu, \nu \in \Omega$, F und G die assoziierten Abbildungen, $\vec{X}_1, \dots, \vec{X}_n$ unabhängige Zufallsvektoren, die die Verteilung ν besitzen; weiter seien

$$\vec{v} = (v_1, \dots, v_n) = F_+((z_{k_1}, \dots, z_{k_n})),$$

$$(z_{k_1}, \dots, z_{k_n}) = S_\pi((\vec{X}_1, \dots, \vec{X}_n)),$$

dann gilt:

a) (zu Lemma 3.4.2)

Die Verteilung von $\vec{v}$ besitzt die Dichte

$$g(\vec{v}) = \begin{cases} |\det \frac{\partial}{\partial \vec{v}} (F_+ \circ G_+^{-1})^{-1}(\vec{v})| n! & \text{falls } \vec{v} \in D_+ \\ 0 & \text{sonst.} \end{cases}$$

b) (zu Satz 3.4.4)

$$|\det \frac{\partial}{\partial \vec{v}} (F_+ \circ G_+^{-1})^{-1}(\vec{v})| = 1 \iff F_+ = G_+.$$

Weiter vergegenwärtige man sich:

$$F_+((z_{k_1}, \dots, z_{k_n})) = (\mu(B_1(\vec{x})), \dots, \mu(B_n(\vec{x})))$$

$$G_+((z_{k_1}, \dots, z_{k_n})) = (\nu(B_1(\vec{x})), \dots, \nu(B_n(\vec{x})))$$

$$\text{wobei } (z_{k_1}, \dots, z_{k_n}) = S_\pi((\vec{x}_1, \dots, \vec{x}_n)) \in S_\pi((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0)$$

ist.

3.5 Definition einer Metrik

Um in nichtparametrischen Testsituationen Gütebetrachtungen anzustellen, ist es nötig, den Hypothesenraum zu metrisieren oder zumindest den einzelnen Elementen einen "Abstand" zur Hypothese zuzuordnen; dies ermöglicht ein etwas weiter gefaßter Abstandsbegriff, der den Hypothesenraum in Klassen einteilt, die durch ihren Abstand zur Hypothese charakterisiert sind. Man erinnere sich beispielsweise an Massey's¹⁾ Berechnung der approximativen Gütefunktion des Kolmogorov - Smirnow - Anpassungstests. In diesem Abschnitt wird eine Metrik konstruiert werden, die Gütebetrachtungen ermöglicht; da sich diese aber nur numerisch in sinnvoller Weise durchführen lassen, wurden sie in der vorliegenden Arbeit nicht ausgeführt. So erhält diese Metrik hier eine rein technische Bedeutung; sie ermöglicht es, den Nachweis zu führen, daß unter noch festzulegenden Voraussetzungen an die Bildung der StEB's aus $F = G$ auch $\mu = v$ gefolgert werden kann.

Zunächst jedoch einige, teilweise als Wiederholung gedachte Definitionen:

- Sei M eine Menge, dann heißt eine Abbildung

$$d: M \times M \longrightarrow \mathbb{R}_+ \cup \{0\}$$

Metrik, wenn sie die folgenden Bedingungen erfüllt:

1) Massey, F. J.: A note on the power of a nonparametric Test. Ann. math. Statistics 20, 440 - 442, (1950).

- i) $(x, y) \in M \times M \implies d(x, y) \geq 0$
- ii) $x, y \in M \implies d(x, y) = d(y, x)$
- iii) $x \in M \implies d(x, x) = 0$
- iv) $(x, y) \in M \times M \implies (d(x, y) = 0 \implies x = y)$
- v) $x, y, z \in M \implies d(x, z) \leq d(x, y) + d(y, z)$

- Unter $\omega(\mathbb{R}_p)$ verstehe man die Menge aller Wahrscheinlichkeitsmaße auf $(\mathbb{R}_p, \mathcal{B}_p)$.

- Ω ist wieder die Menge aller stetigen Wahrscheinlichkeitsmaße mit auf ganz $\mathbb{R}_p$ positiver und stetiger Dichte; $\Omega \subset \omega(\mathbb{R}_p)$.

- F_p sei die Menge aller nach unten unbeschränkten Quader aus $\mathbb{R}_p$.

$$F_p := \{A(\vec{a}) \subset \mathbb{R}_p : -\infty < x_i \leq a_i, i = 1, \dots, p\}$$

- $E \subset \mathcal{B}_p$; E enthalte ein Erzeugendensystem von $\mathcal{B}_p$, welches eine isotone Folge $(E_n: n \in \mathbb{N}) \in \mathcal{B}_p^{\mathbb{N}}$ mit $E_n \uparrow \mathbb{R}_p$ umfaßt.

Unter diesen Voraussetzungen, kann jetzt die folgende Definition getroffen werden:

3.5.1 Definition:

Sei $\mu, \nu \in \omega(\mathbb{R}_p)$ und E eine wie oben beschriebene Teilklasse von $\mathcal{B}_p$.

$$d_0: \omega(\mathbb{R}_p) \times \omega(\mathbb{R}_p) \longrightarrow \mathbb{R}_+ \cup \{0\}$$

$$(\mu; \nu) \longmapsto \sup_{A \in E} |\mu(A) - \nu(A)|.$$

3.5.2. Korollar:

d_0 ist eine Metrik.

Beweis:

Die Eigenschaften i), ii), und iii) einer Metrik sind sofort einzusehen. Es bleibt nur noch iv) und v) zu zeigen.

Zu iv):

$$d_0(\mu; \nu) = 0 \implies \sup_{A \in E} |\mu(A) - \nu(A)| = 0 \implies$$

$$\mu(A) - \nu(A) = 0, \text{ für alle } A \in E \implies \mu(A) = \nu(A)$$

für alle $A \in E$; da E ein Erzeugendensystem von $\mathcal{B}_p$ enthält, folgt hieraus $\mu = \nu$.

Zu v) (Dreiecksungleichung):

$$\mu, \nu, \tau \in \omega(\mathbb{R}_p) \implies$$

$$d_0(\mu; \tau) = \sup_{A \in E} |\mu(A) - \tau(A)|$$

$$\leq \sup_{A \in E} (|\mu(A) - \tau(A)| + |\tau(A) - \nu(A)|)$$

$$= \sup_{A \in E} |\mu(A) - \tau(A)| + \sup_{A \in E} |\tau(A) - \nu(A)|$$

$$= d_0(\mu; \tau) + d_0(\tau; \nu).$$

//.

Genau genommen wurde durch die *Definition 3.5.1* nicht *eine* Metrik festgelegt, sondern eine ganze Klasse von Metriken, die im einzelnen von der zugrunde gelegten Teilmenge von B_p , die den geforderten Regularitätsbedingungen genügt, abhängen. Hier ist schon erkennbar, daß die Partitionenfolgen, die durch StEB's entstehen, solche Teilklassen unter bestimmten Voraussetzungen bilden können. Unter welchen Umständen dies tatsächlich der Fall ist, wird anschließend dargestellt werden. Ausgegangen wird von Ω .

Hierzu sei fest vorgegeben:

- der Stichprobenumfang $n \geq p$
- die Permutation $\pi_n: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$
- die Schnittfunktionen, hier die Projektionen $pr_1, \dots, pr_p$ auf die Koordinatenachsen,
- die Zuordnung der Indices an die Schnittfunktionen: $\pi_n(\{1, \dots, n\}) \longrightarrow \{pr_1, \dots, pr_p\}$; diese Abbildung sei, was keine große Einschränkung darstellt, surjektiv.

Hieraus erhält man, wie in Abschnitt 3.1 gezeigt wurde, aus einem n -Tupel p -dimensionaler Vektoren ein n -Tupel eindimensionaler Größen.

$\vec{x} := (\vec{x}_1, \dots, \vec{x}_n)$ durchlaufe die Menge $(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_{\hat{\pi}}$, somit durchläuft $\vec{z} := (z_{k_1}, \dots, z_{k_n}) = S_{\pi}((\vec{x}_1, \dots, \vec{x}_n))$ die Menge $S_{\pi}(\underbrace{(\mathbb{R}_p \times \dots \times \mathbb{R}_p)_o}_{n\text{-mal}})$; hiermit werden

nun die StEB's $\{B_1(\vec{x}), \dots, B_{n+1}(\vec{x})\}$ konstruiert, die folgender Bedingung zu genügen haben:

$B_1(\vec{x}) \in F_p$ und

$$\{B_1(\vec{x}) : \vec{x} \in (\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0\} = F_p$$

Daß diese Festlegung möglich ist, möge folgendes Beispiel zeigen.

3.5.3 Beispiel:

Sei $\vec{x} = (x_1, \dots, x_n) \in (\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0$, $n \geq p$

$$\pi: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$$

$$k_i = \pi(i) = \begin{cases} p - i + 1 & \text{für } i = 1, \dots, p \\ i & \text{für } i = p + 1, \dots, n \end{cases}$$

$$h_{k_i} = \begin{cases} pr_i & \text{für } i = 1, \dots, p \\ pr_i & \text{für } i = p + 1, \dots, n \end{cases}$$

Unter manchen Gesichtspunkten ist die Wahl $h_{k_i} = pr_i$ für $i = p + 1, \dots, n$ sicher nicht optimal, für die gegenwärtigen Zwecke ist sie jedoch ausreichend.

Mit den in § 1 eingeführten Bezeichnungen ist nun

$$B_1(\vec{x}) = \{\vec{y} \in \mathbb{R}_p : pr_1(\vec{y}) \leq z_{k_1}, \dots, pr_p(\vec{y}) \leq z_{k_p}\} \in F_p.$$

Sei andererseits

$$C := \{\vec{y} \in \mathbb{R}_p : pr_1(\vec{y}) \leq c_1, \dots, pr_p(\vec{y}) \leq c_p\} \in F_p.$$

Es gibt ein $\vec{x}_p \in \mathbb{R}_p$, so daß

$$pr_1(\vec{x}_p) = c_1 \text{ und } \vec{x}_1, \dots, \vec{x}_{p-1}, \vec{x}_{p+1}, \dots, \vec{x}_n, \text{ so daß}$$

$$\text{pr}_1(\vec{x}_1) < \text{pr}_1(\vec{x}_2) < \dots < c_1 < \text{pr}_1(\vec{x}_{p+1}) < \dots < \text{pr}_1(\vec{x}_n),$$

Somit sind die 1. Koordinaten der Elemente $\vec{x}_1, \dots, \vec{x}_n$ bestimmt, hierzu gibt es $\vec{x}_{p-1}$, für das $\text{pr}_2(\vec{x}_{p-1}) = c_2$ gilt

und

$$\text{pr}_2(\vec{x}_1) < \dots < \text{pr}_2(\vec{x}_{p-2}) < c_2 < \text{pr}_2(\vec{x}_p) < \dots < \text{pr}_2(\vec{x}_n),$$

dies ergibt die 2. Koordinaten von $\vec{x}_1, \dots, \vec{x}_n$;
für $2 < i \leq p$ findet man darauf aufbauend die i-ten Koordinaten, so daß

$$\text{pr}_i(\vec{x}_1) < \dots < c_i < \text{pr}_i(\vec{x}_{p-i+2}) < \dots < \text{pr}_i(\vec{x}_n) \text{ gilt.}$$

Hierbei wurde $c_i < \text{pr}_i(\vec{x}_{p-i+2}) < \dots < \text{pr}_i(\vec{x}_n)$ nur festgelegt um sicherzustellen, daß $(\vec{x}_1, \dots, \vec{x}_n) \in (\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0$ wirklich gewährleistet wird; diese Forderungen könnten durch andere, die das gleiche leisten, ersetzt werden.

Bildet man nun aus den so gewonnenen $\vec{x}_1, \dots, \vec{x}_n$ mittels der obigen Festlegungen StEB's, so erhält man

$$B_i(\vec{x}) = C \in \mathbb{F}_p.$$

Es genügt, die Blockbildung jetzt nur noch in Abhängigkeit von $\vec{z} = (z_{k_1}, \dots, z_{k_n}) \in S_{\pi}((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0)$ zu betrachten, dies soll durch die Festlegung

$$B_i(\vec{z}) := B_i(\vec{x}), \quad i = 1, \dots, n+1$$

zum Ausdruck kommen.

Für die weiteren Überlegungen ist es günstiger, nicht mehr die Abbildung F , sondern F_+ und die Vektoren $\vec{v} = (v_1, \dots, v_n) \in D_+$ zu betrachten. Hiermit kann nun eine Metrik d_2 festgelegt werden.

3.5.4 Definition:

Seien $\mu, \nu \in \omega(\mathbb{R}_p)$ und

$$P := \{(B_1(\vec{z}), \dots, B_{n+1}(\vec{z})) :$$

$$\vec{z} \in S_\pi((\mathbb{R}_p \times \dots \times \mathbb{R}_p)_0), B_1(\vec{z}) \in F_p\}$$

$$d_2: \omega(\mathbb{R}_p) \times \omega(\mathbb{R}_p) \longrightarrow \mathbb{R}_+ \cup \{0\}$$

$$d_2(\mu; \nu) := \sup_{A \in P} |\mu(A) - \nu(A)|.$$

3.5.5 Satz:

Sei $\mu, \nu \in \Omega$, $\vec{v} = (v_1, \dots, v_n) \in D_+$,

F_+ die zu μ assoziierte und G_+ die zu ν assoziierte Abbildung, sowie

$$T: D_+ \longrightarrow \mathbb{R}_{n+1}$$

$$T(\vec{v}) := (v_1, \dots, v_n, 1 - \sum_{i=1}^n v_i) =: (v_1, \dots, v_{n+1}) =: \vec{v}_+,$$

so gilt:

$$d_2(\mu, \nu) = \sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |v_i - \text{pr}_i(T(F_+(G_+^{-1}(\vec{v}))))|$$

Beweis:

$$\sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |v_i - \text{pr}_i(T(F_+(G_+^{-1}(\vec{v}))))| =$$

$$\sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |\text{pr}_i(T(G_+(G_+^{-1}(\vec{v})))) - \text{pr}_i(T(F_+(G_+^{-1}(\vec{v}))))| =$$

$$\sup_{\vec{z} \in (T \circ G_+)^{-1}(D_+)} \sup_{i=1, \dots, n+1} |\text{pr}_i(T(G_+(\vec{z}))) - \text{pr}_i(T(F_+(\vec{z})))| =$$

$$\sup_{\vec{z} \in S_\pi((R_p^x \dots R_{p_0}^y))} \sup_{i=1, \dots, n+1} |v(B_i(\vec{z})) - \mu(B_i(\vec{z}))| =$$

$$\sup_{A \in \mathcal{P}} |v(A) - \mu(A)| = d_2(v; \mu) \quad //.$$

Dieser Satz ermöglicht es nun, unter den obigen Voraussetzungen an Ω und $\mathcal{E}$ (d.h. an die Bildung der StEB's) aus der Gleichheit der Abbildungen F_+ und G_+ auf die Gleichheit der Wahrscheinlichkeitsmaße μ und ν zu schließen.

3.5.6 Satz:

Seien $\mu, \nu \in \Omega$ und F und G die assoziierten Abbildungen, sowie d_2 die in 3.5.4 definierte Metrik, dann gilt mit Satz 3.5.5

$$F = G \iff \mu = \nu. \quad 1)$$

1) F_+ und G_+ sind durch F und G eindeutig bestimmt.

Beweis:

" $\Leftarrow$ ": Diese Richtung ist sofort einzusehen.

" $\Rightarrow$ ":

$$\begin{aligned} d_2(\mu; \nu) &= \sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |v_i - \text{pr}_i(T(F_+(C_+^{-1}(\vec{v}))))| \\ &= \sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |v_i - \text{pr}_i(T(F_+(F_+^{-1}(\vec{v}))))| \\ &= \sup_{\vec{v} \in D_+} \sup_{i=1, \dots, n+1} |v_i - v_i| = 0 \quad \Rightarrow \\ \mu &= \nu \end{aligned}$$

//

Diesen Paragraphen zusammenfassend konnt¹⁾ gezeigt werden: ist μ ein Wahrscheinlichkeitsmaß aus Ω und wurden die StEB's so gebildet, daß sie ein Element aus F_p enthalten, so ist der durch Integration dieser Mengen bzgl. μ entstandene Zufallsvektor $(V_1, \dots, V_n)$ genau dann *Dirichlet - verteilt*¹⁾, wenn μ das der Stichprobe $\vec{X}_1, \dots, \vec{X}_n$ zugrunde liegende Wahrscheinlichkeitsmaß war. Gegenstand des nächsten Paragraphen wird sein, einen Test, dessen Testgröße auf dem Vektor $(V_1, \dots, V_n)$ aufbaut, zu folgender Situation anzugeben:

$H_0: \mu = \mu_0, \mu_0 \in \Omega$ gegen

$H_1: \mu \neq \mu_0, \mu \in \Omega.$

T. W. Anderson²⁾ hat in seinem Beitrag gezeigt, daß unter bestimmten Voraussetzungen an die Bildung von StEB's es möglich ist, für wachsenden Stichprobenumfang n eine "falsche" Verteilungsannahme μ zu erkennen; wie schon in Paragraph zwei erwähnt wurde, hat er diese Aussage jedoch nicht im strengen Sinn bewiesen.

Sein Vorgehen soll als Beispiel noch einmal aufgegriffen werden.

1) siehe S. 95

2)

3.5.7 Beispiel:

Sei Ω der Hypothesenraum, $\mu, \nu \in \Omega$

Die Elemente der Blockbildung seien folgendermaßen definiert

- $n = 2^q - 1$, $q \in \mathbb{N} \cup \{0\}$, n ist der Stichprobenumfang

- $\{pr_1, pr_2\}$ die Menge der Schnittfunktionen

- $\pi_n: \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ die Permutation

$$k_1(n) := \lfloor n \cdot \frac{1}{2} \rfloor + 1$$

$$k_2(n) := \lfloor n \cdot \frac{1}{4} \rfloor + 1$$

$$k_3(n) := \lfloor n \cdot \frac{3}{4} \rfloor + 1$$

allgemein:

$$k_i(n) = \lfloor n \cdot \frac{1}{2^{l+1}} ((i-2^l) \cdot 2 + 1) \rfloor + 1$$

falls $2^l \leq i < 2^{l+1}$, $l = 0, \dots, q$.

- zugeordnet wird

$$k_1(n) \longrightarrow pr_1$$

$$k_2(n) \longrightarrow pr_2$$

$$k_i(n) \longrightarrow \begin{cases} pr_1 & \text{falls } 2^l \leq i < 2^{l+1} \text{ und } l \text{ gerade} \\ & \text{oder } l = 0 \text{ gilt} \\ pr_2 & \text{falls } 2^l \leq i < 2^{l+1} \text{ und } l \text{ ungerade ist} \end{cases}$$

Aufgrund dieser Maßgaben ist aus einer Stichprobe unabhängiger und identisch verteilter Zufallsvektoren $\vec{X}_1, \dots, \vec{X}_n$, für alle $n = 2^q - 1$, $q \in \mathbb{N} \cup \{0\}$ der Vektor $(Z_{k_1(n)}, \dots, Z_{k_n(n)})$ zu gewinnen, der wiederum die StEB's festlegt, sowie die Abbildungen ${}_nF$ bzgl. μ und ${}_nG$ bzgl. ν , wobei n jetzt den Stichprobenumfang angibt. ${}_nF_k$, ${}_nG_k$ bezeichnen die jeweiligen partiellen Abbildungen.

Sei weiter

$$\vec{e}_n := \left(\frac{1}{2}, \frac{1}{4}, \dots, \frac{2^q - 1}{2^q} \right), \quad n = 2^q - 1$$

und

$$\vec{b}_n = (b_{k_1(n)}, \dots, b_{k_n(n)}) := {}_nG^{-1}(\vec{e}_n)$$

$$\vec{a}_n = (a_{k_1(n)}, \dots, a_{k_n(n)}) := {}_nF^{-1}(\vec{e}_n)$$

$\vec{a}_n$ und $\vec{b}_n$ bestimmen jeweils Partitionen von $\mathbb{R}_2$, deren Elemente bzgl. μ bzw. ν die Masse $\frac{1}{2^q}$ besitzen und Quader sind. Sie bilden somit die Partition, die auch T. W. Anderson in seinen Konsistenzbetrachtungen konstruiert hatte¹⁾.

Ist ν die zugrunde liegende Verteilung und werden die StEB's mit $\mu * \nu$ integriert, so strebt die Wahrscheinlichkeit, daß die durch das Verfahren gewonnenen StEB's mit wachsendem n (bzw. q) sich von den vorgegebenen Quadern (bestimmt durch ${}_nF^{-1}(\vec{e}_n)$) unterscheiden, gegen 1. Nachgewiesen wird diese Aussage, indem gezeigt wird, daß $(Z_{k_1(m)}, \dots, Z_{k_n(m)})$ mit $m = 2^k - 1$, $k > q$ für $m \rightarrow \infty$ schwach gegen $\vec{b}_n * \vec{a}_n$ konvergiert:

¹⁾ siehe Seite 7 unter 2)

Es gilt:

${}_m G_n((U_{k_1(m)}, \dots, U_{k_n(m)})) = (U_{k_1(m)}, \dots, U_{k_n(m)})$ und

$U_{k_i(m)}$ ist verteilt wie die $k_i(m)$ -te Orderstatistik von m unabhängigen und identisch auf $[0;1]$ gleichverteilten Zufallsvariablen, wegen der speziellen Wahl der $k_i(n)$ konvergiert $U_{k_i(m)}$ gegen das

$\frac{(i-2^{l+1})2+1}{2^{l+1}}$ Quantil, $2^l \leq i < 2^{l+1}$, $l = 0, \dots, q$ und

somit gegen $\frac{(i-2^{l+1})2+1}{2^{l+1}}$; daraus folgt

$(U_{k_1(m)}, \dots, U_{k_n(m)})$ konvergiert stochastisch gegen

$\vec{e}_n$ und wegen der Stetigkeit von ${}_m G_n^{-1}$ konvergiert

$(Z_{k_1(m)}, \dots, Z_{k_n(m)})$ von der Nummer n an stochastisch

gegen $\vec{b}_n$. Nach T. W. Anderson¹⁾ gibt es einen Index, wieder mit n benannt, so daß $\vec{b}_n \neq \vec{a}_n$ gilt. Das bedeutet, daß sich die durch fortgesetzte Zweiteilung bezüglich μ konstruierten Quader von denen, die durch fortgesetzte Zweiteilung bezüglich ν entstanden sind, unterscheiden.

1) Siehe Seite 7 unter 2)

§ 4

Vorschlag einer Testgröße

Ausgegangen wird auch in diesem Paragraphen von einer Stichprobe n unabhängiger, p -dimensionaler $p \geq 2$, identisch verteilter Zufallsvektoren, deren Verteilung μ eine stetige und auf ganz $\mathbb{R}_p$ positive Dichte besitze; der entsprechende Hypothesenraum sei wieder mit Ω bezeichnet. Zugrunde liege nun die Testsituation:

$$H_0: \mu = \mu_0$$

$$H_1: \mu \neq \mu_0, \mu \in \Omega \setminus \{\mu_0\}.$$

Bisher wurde gezeigt:

Bildet man aus der Stichprobe $\hat{X}_1, \dots, \hat{X}_n$ die StEB's und integriert diese mit μ_0 , so sind die daraus entstandenen Zufallsvariablen $V_1, \dots, V_n$ bzw. $U_1, \dots, U_n$ dann und nur dann *Dirichlet*-verteilt bzw. verteilt wie die Orderstatistik n unabhängiger und identisch auf $[0;1]$ gleichverteilter Zufallsgrößen, wenn μ_0 die *tatsächlich* zugrunde liegende Verteilung war.

Das Problem reduziert sich somit darauf, einen Anpassungstest für eine der beiden aus μ_0 hervorgegangenen Hypothesen zu finden. In der Literatur finden sich verschiedene Tests, die dieser Prüfsituation angemessen sind; am bekanntesten sind wohl der *Kolmogorov-Smirnov-Test* und der Test von *Cramer-von Mises-Smirnov*, deren Testgrößen beide auf einer Differenz zwischen

der Verteilung der Hypothese und der, aus der Stichprobe gewonnenen, empirischen Verteilungsfunktion bestehen. Im vorliegenden Fall wären sie sicher angebracht, wenn beabsichtigt wäre, einen Test, der auf dem Zufallsvektor $(U_1, \dots, U_n)$ basiert, zu konstruieren. Für die zugrunde liegende Situation erscheint es jedoch günstiger, wie im folgenden kurz begründet werden wird, sich bei der Suche nach einem geeigneten

Test auf die Zufallsgrößen $V_1, \dots, V_n$, $1 - \sum_{i=1}^n V_i$ zu stützen. Im eindimensionalen Fall ist die Orderstatistik nach der Beobachtung der Stichprobe eindeutig festgelegt; nicht so bei einer Stichprobe, die nach der Bildung von StEB's entstanden ist:

Auch wenn die Schnittfunktionen, die Permutation und die Zuordnung der Indices an die Schnittfunktionen feststeht, ist für eine beliebige Permutation $j_1, \dots, j_n$ von $1, \dots, n$ der Zufallsvektor $(U'_1, \dots, U'_n)$

$$U'_1 := V_{j_1}, \quad U'_2 := V_{j_1} + V_{j_2} \quad \text{usw.}$$

ebenfalls verteilt wie die Orderstatistik n unabhängiger, identisch auf $[0;1]$ gleichverteilter Zufallsvariabler. Dies gilt nicht für den folgenden Zufallsvektor $(U_{(1)}, \dots, U_{(n)})$, der wie folgt definiert ist:

$V_{(1)}, \dots, V_{(n)}$ sei die geordnete Stichprobe von $V_1, \dots, V_n$ und $U_{(1)} := V_{(1)}$ sowie $U_{(2)} := V_{(1)} + V_{(2)}$ usw.

Der Zufallsvektor $(V_{(1)}, \dots, V_{(n)})$ und somit $(U_{(1)}, \dots, U_{(n)})$ schafft aber eine weitere Möglichkeit, eine vorliegende Verteilung $\mu \neq \mu_0$ zu erkennen; die in diesen Zufallsgrößen liegende Information wird aber von den Tests von Kolmogorov - Smirnov und Cramer - von Mises - Smirnov nicht verarbeitet.

Zur Motivierung der einzuführenden Testgröße mögen die folgenden Überlegungen dienen.

$V_1, \dots, V_n$ haben, wenn die Hypothese H_0 zutrifft, den

gleichen Erwartungswert $1/(n+1)$. Dies ist heuristisch darauf zurück zu führen, daß im Fall der Nullhypothese μ_0 der Stichprobenraum IR_p in $n+1$, bzgl. des Maßes μ_0 "gleich große" Teilmengen (die StEB's) aufgeteilt wird. Darüber hinaus sind die $V_1, \dots, V_n$ durch die Wahl des speziellen Verfahrens *eindeutig* bestimmt. Aus diesen Gründen wird eine auf dem Zufallsvektor $(V_1, \dots, V_n)$ basierende Teststatistik gewählt werden. Wie oben schon begründet wurde, liegt es nahe, als Testgröße eine Funktion der Differenzen $V_i - 1/(n+1)$ oder eine Funktion der Orderstatistik zu betrachten; im vorliegenden Fall fiel die Entscheidung auf eine Teststatistik der Gestalt

$$\sum_{i=1}^n |V_i - 1/(n+1)| + |1 - \sum_{i=1}^n V_i - 1/(n+1)|^1,$$

da hier die Abweichungen von $1/(n+1)$, die bei einer Verteilung der Alternative größer sein werden, durch die Summation noch einmal verdeutlicht werden. In dieser Hinsicht gleichrangig dürfte auch die Funktion

$$\sum_{i=1}^n (V_i - 1/(n+1))^2 + (1 - \sum_{i=1}^n V_i - 1/(n+1))^2 \quad 2)$$

sein; sie erschien aber als Testgröße, abgesehen von technischen Schwierigkeiten, auch in Hinblick auf eine Abschätzung der Gütefunktion als ungünstiger. Die Verteilung beider Teststatistiken ist in der Literatur^{1, 2)} bekannt; jedoch ist anzumerken, daß von der

1) Sherman, B.: A random variable related to the Spacing of Sample values. Ann. of math. Statistics 21, 339 - 341 (1950).

2) Kimball, B. F.: Some basic theorems for developing Tests of Fit for the case of the nonparametric probability distribution function I. Ann. of math. Statistics 13, 540 - 548 (1947).

zweiten Testgröße nur die Momente und die charakteristische Funktion angegeben wurde. Bevor jedoch auf die Testgröße

$\sum_{i=1}^n |V_i - 1/(n+1)| + |1 - \sum_{i=1}^n V_i - 1/(n+1)|$ eingegangen wird, sollen im folgenden Abschnitt noch einige Eigenschaften der *Dirichlet* - Verteilung angeführt werden.

4.1 Bemerkungen zur Dirichlet - Verteilung

In diesem Abschnitt werden einige Eigenschaften der *Dirichlet* - Verteilung referiert, wobei der Darstellung, wie sie N. L. Johnson und S. Kotz in ihrem Buch "Distributions in Statistics, Continuous Multivariate Distributions" (Wiley series in Probability and math. Statistics. New York, London, Sydney, Toronto (1972)) gefolgt wird. Auf Beweise und Herleitungen wird in diesem Abschnitt verzichtet.

Ausgegangen wird von $m+1$ unabhängigen Zufallsvariablen $X_0, X_1, \dots, X_m$, die χ^2 -verteilt sind mit den - nicht notwendig identischen - Freiheitsgraden $v_0, \dots, v_m$, $v_i \in \mathbb{R}_+$, $i=0, \dots, m$;
daraus werden die m Zufallsvariablen

$Y_j := X_j \cdot \left(\sum_{i=0}^m X_i \right)^{-1}$, $j=1, \dots, m$ gebildet.

Die Verteilung des Zufallsvektors $(Y_1, \dots, Y_m)$ ist dann die *Dirichlet* - Verteilung mit den Parametern $\theta_0, \dots, \theta_m$, wobei $\theta_i := 1/2 \cdot v_i$, $i=0, \dots, m$ gilt.

Die Dichte dieser Verteilung hat die Gestalt:

$$f((y_1, \dots, y_m)) = \frac{\Gamma(\sum_{j=1}^m \theta_j)}{\prod_{j=0}^m \Gamma(\theta_j)} (1 - \sum_{j=1}^m y_j)^{\theta_0 - 1} \prod_{j=1}^m y_j^{\theta_j - 1}$$

falls $0 \leq y_j$, $j=1, \dots, m$ und $\sum_{j=1}^m y_j \leq 1$ gilt; ansonsten ist $f((y_1, \dots, y_m)) = 0$.

Die eindimensionalen Randverteilungen von Y_k sind *Beta - Verteilungen* mit jeweils den Parametern

$$\theta_k, \sum_{j=0}^m \theta_j - \theta_k, \quad \text{aus diesem Grunde wird die}$$

Dirichlet - Verteilung auch als mehrdimensionale Verallgemeinerung der *Beta - Verteilung* angesprochen.

Sei $\Theta := \sum_{i=0}^m \theta_i$, so sind die s -dimensionalen Randver-

teilungen wieder *Dirichlet - Verteilungen* mit den Parametern $\theta_1, \dots, \theta_s$, $\Theta - \sum_{i=1}^s \theta_i$.

Ist $(Y_1, \dots, Y_m)$ *Dirichlet - verteilt* mit den Parametern $\theta_0, \dots, \theta_m$ so gilt:

$$E(Y_i) = \frac{\theta_i}{\Theta} \quad i=1, \dots, m$$

$$V(Y_i) = \frac{\theta_i (\Theta - \theta_i)}{\Theta^2 (\Theta + 1)} \quad i=1, \dots, m.$$

Die Kovarianzen haben die Werte:

$$\text{Cov}(Y_i; Y_j) = - \frac{\theta_i \theta_j}{\Theta^2 (\Theta + 1)} \quad i \neq j, \quad i, j=1, \dots, m$$

Die Verteilung des Zufallsvektors $(V_1, \dots, V_m)$, der durch die Bildung von StEB's entstanden ist, ist also eine *Dirichlet - Verteilung* mit den Parametern $\theta_i = 1$, $i=0, \dots, m$. Die entsprechenden Erwartungswerte haben folglich den Wert $1/(m+1)$.

4.2 Die Testgröße

Zu Grunde liege die folgende Testsituation:

- $\vec{X}_1, \dots, \vec{X}_n$ seien n unabhängige und identisch verteilte, stetige Zufallsvektoren auf $(\mathbb{R}_p; \mathcal{B}_p)$;
- Ω bezeichne die Menge aller stetigen Wahrscheinlichkeitsmaße mit stetiger, auf ganz $\mathbb{R}_p$ positiver Dichte; Ω bezeichne den Hypothesenraum.
- Basierend auf einer Realisation $\vec{x}_1, \dots, \vec{x}_n$ dieser Stichprobe soll die Hypothese
 $H_0: \mu = \mu_0$ gegen die Alternative
 $H_1: \mu \neq \mu_0, \mu \in \Omega \setminus \{\mu_0\}$
 getestet werden.
- Das Signifikanzniveau sei $\alpha \in (0; 1)$.

In bekannter Weise wird aus $(\vec{x}_1, \dots, \vec{x}_n)$ durch die Bildung von StEB's und durch Integration der Blöcke mit dem hypothetischen Wahrscheinlichkeitsmaß

der Vektor $(v_1, \dots, v_n, 1 - \sum_{i=1}^n v_i) := (v_1, \dots, v_{n+1})$

erhalten. Mit diesem Vektor soll dann die Hypothese $H'_0: (v_1, \dots, v_{n+1})$ ist die Realisation eines *Dirichlet*-verteilten Zufallsvektors mit den Freiheitsgraden $\theta_0 = 1, \dots, \theta_n = 1$ - zum Niveau $\alpha \in (0; 1)$ überprüft werden. Hierbei ist der Hypothesenraum die Menge aller durch die Transformationen $F \circ G^{-1}$ hervorgegangenen Verteilungen. F bezeichne die zu μ_0 assoziierte Abbildung und G die zu einem $\mu \in \Omega$ assoziierte Abbildung.

Wie schon im vorangegangenen Abschnitt angekündigt, ist die Testgröße

- $S((V_1, \dots, V_{n+1})) := 1/2 \sum_{i=1}^{n+1} |V_i - 1/(n+1)|$, wobei der Faktor $1/2$ nur rechentechnische Bedeutung hat.
- τ_α Sei das α -Quantil der Verteilung von $S((V_1, \dots, V_{n+1}))$, falls der Zufallsvektor $(V_1, \dots, V_n)$ *Dirichlet*-verteilt ist mit den Parametern $\theta_0 = 1, \dots, \theta_n = 1$.

Der Test ϕ hat dann die Gestalt:

$$\phi((\vec{X}_1, \dots, \vec{X}_n)) = \begin{cases} 1 & \text{falls } S((V_1, \dots, V_{n+1})) \geq \tau_\alpha \\ 0 & \text{sonst} \end{cases}.$$

4.3 Die Verteilung der Testgröße

B. Sherman¹⁾ gibt in seiner Arbeit die Verteilungsfunktion der Testgröße $S((V_1, \dots, V_{n+1}))$ an, die aus den Momenten der Verteilung und darauf aufbauend aus der charakteristischen Funktion dieser Verteilung zu erhalten ist. Da er in diesem Artikel jedoch nur die Berechnung der Momente ausführt, wird am Ende dieses Abschnittes eine Herleitung der Dichte (der Verteilungsfunktion) dieser Verteilung skizziert. Zunächst sei jedoch:

$X := S((V_1, \dots, V_{n+1}))$, wobei der Vektor $(V_1, \dots, V_n)$ *Dirichlet*-verteilt mit den Parametern $\theta_0 = 1, \dots, \theta_n = 1$ ist; die Verteilungsfunktion $F(x)$ hat dann die Gestalt:

$$F(x) = 1 + \sum_{q=0}^{n-r-1} \sum_{p=0}^q (-1)^{q-p+1} \binom{q}{p} \binom{n+1}{q+1} \binom{n+q-p}{q} \binom{n-1}{n+1} \binom{n-q}{n+1} (n-q-x)^{n-p}$$

¹⁾ siehe Seite 94 unter 1)

für $x \in [\frac{r}{n+1} ; \frac{r+1}{n+1})$ und $r=0, \dots, n-1$;

für $x < 0$ gilt $F(x) = 0$ und

für $\frac{n}{n+1} \leq x < \infty$ gilt $F(x) = 1$.

Wie man sofort sieht, ist $F(x)$ in jedem Intervall

$[\frac{r}{n+1} ; \frac{r+1}{n+1})$, $r=0, \dots, n-1$, ein Polynom n -ten Grades.

Aus dieser Verteilungsfunktion muß dann, zur Durchführung des Tests das Quantil τ_α bestimmt werden;

τ_α ist der Wert, der die Gleichung

$$F(\tau_\alpha) = 1 - \alpha$$

löst.

Ist einmal der Bereich $[\frac{l}{n+1} ; \frac{l+1}{n+1})$, $l \in \{0, \dots, n-1\}$

gefunden, der τ_α umfaßt, so ist zur endgültigen Bestimmung des Tests ϕ das Polynom

$$P(x) = \alpha + \sum_{q=0}^{n-l-1} (-1)^{q-p+1} \binom{q}{p} \binom{n+1}{q+1} \binom{n+q-p}{q} \left(\frac{n-q}{n+1}\right)^p \left(\frac{n-q}{n+1} - x\right)^{n-p}$$

zu radizieren; eine Aufgabe, die bei den zu erwartenden großen Stichprobenumfängen n wohl nur mit numerischen Methoden zu lösen ist.

Herleitung der Dichte der Verteilung der Testgröße:

Die Dichte dieser Verteilung wird mit Hilfe der charakteristischen Funktion berechnet, die man zunächst in der, in diesem Fall auf ganz $\mathbb{R}$ konvergenten Reihenentwicklung, darstellt und umformt. Zu dieser Reihenentwicklung benötigt man die von Sherman¹⁾ berechneten Momente der Verteilung.

1)
Siehe Seite 94 unter 1).

Das k -te Moment dieser Verteilung lautet:

$$a_{n,k} = \binom{n+k}{k}^{-1} \sum_{s=0}^k \binom{n+1}{s+1} \binom{k-1}{s} \left(\frac{n-s}{n+1}\right)^{n+k}$$

n bezeichne den Stichprobenumfang und

$$k = \begin{cases} k-1 & \text{falls } k \leq n \\ n-1 & \text{falls } k > n \end{cases}$$

Zunächst zwei Bezeichnungen:

- $F^n(g(x))$ bezeichne ein " n -faches unbestimmtes Integral" der Funktion g , falls es existiert.¹⁾
- Sei f die Dichte einer Verteilung, dann wird die charakteristische Funktion mit
$$\phi(t) := \int f(x) \cdot e^{itx} dx$$
 bezeichnet.

Um die Dichte dieser Verteilung zu berechnen, wird zunächst die Fourier-Transformierte in ihrer Reihenentwicklung dargestellt und diese dann durch Bildung von höheren Ableitungen und mehrfacher Bildung von Stammfunktionen äquivalent umgeformt bis diese Reihe als *endliche* Summe in Funktionen von Ableitungen von Exponentialfunktionen dargestellt ist. Mit dieser Form kann dann die Dichte berechnet werden, indem die Inversionsformel auf die ermittelten Summanden angewendet wird. Dies führt dann zu Funktionen (die mit den gegenwärtig verwendeten Beziehungen im Anhang zu § 4 aufgeführt sind), die durch ihr Verhalten für unendlich große Werte zum einen die Fallunterscheidungen in der Darstellung der Dichte "automatisch" bedingen und zum anderen zur Dichte selbst führen. Aus dieser Dichte wird abschließend die Verteilungsfunktion berechnet.

1) siehe Seite VI

Zur Umformung der Reihenentwicklung von $\phi(t)$:

$$\phi(t) = \left(\sum_{k=1}^n \sum_{s=0}^{k-1} \binom{n+k}{k}^{-1} \binom{n+1}{s+1} \binom{k-1}{s} \left(\frac{n-s}{n+1} \right)^{n+k} \frac{1}{k!} (it)^k + \right. \\ \left. \sum_{k=n+1}^{\infty} \sum_{s=0}^{n-1} \binom{n+k}{k}^{-1} \binom{n+1}{s+1} \binom{k-1}{s} \left(\frac{n-s}{n+1} \right)^{n+k} \frac{1}{k!} (it)^k \right) + 1$$

Die folgende Umformung beruht im ersten Schritt auf einer Vertauschung der Summationsindices der Doppelsummen. Bei der ersten Doppelsumme ist dies das übliche übliche Vorgehen, wie es nachstehend im Schema aufgeführt wird:

$$\sum_{k=1}^n \sum_{s=0}^{k-1} b_{k,s} = \sum_{s=0}^{n-1} \sum_{k=s+1}^n b_{k,s}.$$

Da bei der zweiten Doppelsumme die Summationsindices nicht voneinander abhängen, ist hier die Vertauschung unproblematisch.

Im zweiten Schritt werden die so erhaltenen Summen zusammengefaßt; schematisiert stellt sich die Summe dann wie folgt dar:

$$\phi(t) = 1 + \sum_{s=0}^{n-1} \sum_{k=s+1}^n b_{k,s} + \sum_{s=0}^{n-1} \sum_{k=n+1}^{\infty} b_{k,s} = \\ 1 + \sum_{s=0}^{n-1} \left(\sum_{k=s+1}^n b_{k,s} + \sum_{k=n+1}^{\infty} b_{k,s} \right) = \\ 1 + \sum_{s=0}^{n-1} \sum_{k=s+1}^{\infty} b_{k,s}$$

Nach diesen Umformungen hat die Formel eine Gestalt, die es erlaubt, die Inverse der *Fourier - Transformatierten* relativ einfach zu berechnen.

Schreibt man diese Formel aus, so ergibt sich:

$$\begin{aligned}\phi(t) &= 1 + \sum_{s=0}^{n-1} \sum_{k=s+1}^{\infty} \binom{n+k}{n}^{-1} \binom{n+1}{s+1} \binom{k-1}{s} \left(\frac{n-s}{n+1}\right)^{n+k} \frac{1}{k!} (it)^k = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} \left(\frac{n-s}{n+1}\right)^n\right) t^{s+1} \cdot \sum_{k=s+1}^{\infty} \frac{(k-1)!}{(n+k)! (k-1-s)!} i^k t^{k-s-1} \left(\frac{n-s}{n+1}\right)^k = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} \left(\frac{n-s}{n+1}\right)^n\right) t^{s+1} \cdot \frac{d^s}{dt^s} \left(\sum_{l=0}^{\infty} \left(\frac{n-s}{n+1}\right)^{l+1} \frac{i^{l+1}}{(n+1+l)!} t^{l+1} \right) = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} \left(\frac{n-s}{n+1}\right)^n\right) t^{s+1} \cdot \frac{d^s}{dt^s} \left(t^{-n-1} \sum_{l=0}^{\infty} \left(\frac{n-s}{n+1}\right)^{l+1} \frac{i^{l+1}}{(n+1+l)!} t^{l+1+n+1} \right) = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} \left(\frac{n-s}{n+1}\right)^n\right) t^{s+1} \cdot \frac{d^s}{dt^s} \left(t^{-n-1} \cdot F \left(\left(\frac{n-s}{n+1} \right)^{l+1} \frac{i^{l+1}}{l!} t^{l+1} \right) \right) = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} \left(\frac{n-s}{n+1}\right)^n\right) t^{s+1} \cdot i \frac{d^s}{dt^s} t^{-n-1} F \left(e^{i((n-s)/(n+1))t} \right) = \\ &= 1 + \sum_{s=0}^{n-1} \left(\frac{n+1}{s+1} \frac{n!}{s!} i^{-n} t^{s+1}\right) \cdot \frac{d^s}{dt^s} t^{-n-1} e^{i((n-s)/(n+1))t}.\end{aligned}$$

Zur Berechnung von $\frac{d^s}{dt^s} t^{-n-1} e^{i((n-s)/(n+1))t}$

verwendet man die *Leibniz'sche*¹⁾ Formel; dies ergibt:

$$\frac{d^s}{dt^s} t^{-n-1} e^{i((n-s)/(n+1))t} =$$

$$\sum_{j=0}^s \binom{s}{j} \frac{(n+j)!}{n!} (-1)^j \left(i \left(\frac{n-s}{n+1}\right)\right)^{s-j} t^{-n-j-1} e^{i((n-s)/(n+1))t}.$$

1) Siehe Anhang Seite 113.

Die charakteristische Funktion hat somit die folgende Gestalt:

$$\phi(t) = 1 + \sum_{s=0}^{n-1} \frac{s}{j} \binom{n+1}{s+1} \binom{s}{j} \frac{(n+j)!}{s!} (-1)^j j!^{-n+s-j} \cdot$$

$$\left(\frac{n-s}{n+1}\right)^{s-j} t^{s-n-j} e^{i((n-s)/(n+1))t}.$$

Nach der *Inversionsformel für Fourier - Transformierte* läßt sich die Dichte f aus ihrer charakteristischen Funktion wieder berechnen:

$$f(x) = \lim_{N \rightarrow \infty} \frac{1}{2\pi} \int_{-N}^N \phi(t) e^{-itx} dt$$

Da im vorliegenden Fall $\phi(t)$ eine Summe in den Funktionen $(t^{s-n-j} e^{i((n-s)/(n+1))t})$ und 1 ist und weiter

$$\lim_{N \rightarrow \infty} \frac{1}{2\pi} \cdot \int_{-N}^{+N} e^{-itx} dt = 0$$

gilt, genügt es, zunächst das Integral von $t^{s-n-j} e^{it((n-s)/(n+1)-x)}$ und somit die Summanden der Dichte zu berechnen.

$$\lim_{N \rightarrow \infty} \frac{1}{2\pi} \int_{-N}^{+N} \frac{1}{t^{n+j-s}} e^{it((n-s)/(n+1)-x)} dt = 1)$$

$$\lim_{N \rightarrow \infty} \frac{1}{2\pi} \left(-e^{it((n-s)/(n+1)-x)} \sum_{k=1}^{n+j-s-1} \frac{(i((n-s)/(n+1)-x))^k}{(n+j-s-1) \cdots (n+j-s-k)} t^{n+j-s-k} + \right.$$

$$\left. \frac{(i((n-s)/(n+1)-x))^{n+j-s-1}}{(n+j-s-1)!} \cdot \text{Ei}(it((n-s)/(n+1)-x)) \right) \Big|_{-N}^{+N} = 2)$$

1) I. S. Gradshteyn, I. M. Ryzhik: Table of Integrals, Series, and Products, corrected and enlarged edition. Acad. Press New York, London, Toronto, Sydney, San Francisco, p. 92, (1980).

2) ... Exponential-Integral-Funktion

$$\lim_{M \rightarrow -\infty} \frac{1}{2\pi} \frac{(i((n-s)/(n+1)-x))^{n+j-s-1}}{(n+j-s-1)!}.$$

$$\begin{aligned} & (ci((\frac{n-s}{n+1}-x)M) - i \cdot si((\frac{n-s}{n+1}-x)M) - ci((\frac{n-s}{n+1}-x)M) - i \cdot si((\frac{n-s}{n+1}-x)M)) = \\ & - \frac{1}{\pi} i \lim_{M \rightarrow -\infty} si((\frac{n-s}{n+1}-x)M) \frac{(i((n-s)/(n+1)-x))^{n+j-s-1}}{(n+j-s-1)!} \quad 2) \end{aligned}$$

$$\lim_{M \rightarrow -\infty} si((\frac{n-s}{n+1}-x)M) = \begin{cases} 0 & \text{für } \frac{n-s}{n+1} - x < 0 \\ -\pi & \text{für } \frac{n-s}{n+1} - x > 0 \end{cases}$$

Da die betrachtete Zufallsvariable X per definitionem nichtnegativ ist, ist es nötig, die obere Grenze ihres Definitionsbereichs zu ermitteln. Zu diesem Zweck ist der Grenzübergang $M \rightarrow -\infty$ besser geeignet. Aus den obigen Fallunterscheidungen folgt nämlich direkt:

$$f(x) = 0 \quad \text{für } n/(n+1).$$

Weiter ist daran abzulesen:

$$\text{für } \frac{r}{n+1} < x < \frac{r+1}{n+1} \quad r=0, \dots, n-1 \text{ stellen nur}$$

die Summanden zum Index $s \leq n-r-1$ einen Beitrag zur Dichte f . Dies ergibt für $f(x)$ in diesem Intervall:

$$\begin{aligned} f(x) &= \sum_{s=0}^{n-r-1} \sum_{j=0}^s \binom{n+1}{s+1} \binom{s}{j} \binom{n+j}{s} (n+j-s) \cdot (-1)^j \\ & \quad \cdot \frac{(n+s-j)}{(n+s-j)} \left(\frac{n-s}{n+1}\right)^{s-j} \left(\frac{n-s}{n+1} - x\right)^{n+j-s-1} \end{aligned}$$

1) si bezeichne die Sinus-Integral-Funktion und ci die Kosinus-Integral-Funktion.

2) siehe Anhang, Seite 112

Wegen der Stetigkeit von $f(x)$ in diesem Intervall kann o.E. der Randpunkt $\frac{r}{n+1}$ einbezogen werden.

für $x \leq 0$ und $x > \frac{n}{n+1}$ ist $f(x) = 0$.

Setzt man noch $s = q$ und $s - j = p$, so ergibt

sich für $x \in [\frac{r}{n+1}; \frac{r+1}{n+1})$ und $r=0, \dots, n-1$:

$$f(x) = \sum_{q=0}^{n-r-1} \sum_{p=0}^q \binom{n+1}{q+1} \binom{q}{p} \binom{n+q-p}{q} (-1)^{q-p} (n-p) \left(\frac{n-q}{n+1}\right)^p \left(\frac{n-q}{n+1} - x\right)^{n-p-1} \quad 1)$$

Zur Berechnung der Verteilungsfunktion $F(x)$ sei

$$x \in [\frac{r}{n+1}; \frac{r+1}{n+1}) , \quad r = 0, \dots, n-1 .$$

Die Stammfunktion der Dichte in diesem Intervall (ohne Konstante) lautet:

$$F_r(x) := \sum_{q=0}^{n-r-1} \sum_{p=0}^q \binom{n+1}{q+1} \binom{q}{p} \binom{n+q-p}{q} (-1)^{q-p} \left(\frac{n-q}{n+1}\right)^p \left(\frac{n-q}{n+1} - x\right)^{n-p} (-1) .$$

Somit ist

$$F(x) = 1 - \left(\sum_{k=r+1}^{n-1} \left(F_k\left(\frac{k+1}{n+1}\right) - F_k\left(\frac{k}{n+1}\right) \right) + \left(F_r\left(\frac{r+1}{n+1}\right) - F_r(x) \right) \right) =$$

$$1 - \left(\sum_{k=r}^{n-2} \left(F_k\left(\frac{k+1}{n+1}\right) - F_{k+1}\left(\frac{k+1}{n+1}\right) \right) + \left(F_{n-1}\left(\frac{n}{n+1}\right) - F_r(x) \right) \right)$$

1) B. Sherman gibt in seinem Artikel anstelle des Binomialkoeffizienten $\binom{n+q-p}{q}$ den Koeffizienten $\binom{n+q-p}{n}$ an.

es gilt:

$$\alpha) F_{n-1}\left(\frac{n}{n+1}\right) = 0 \quad \text{und}$$

$$\beta) F_{k\left(\frac{k+1}{n+1}\right)} - F_{k+1}\left(\frac{k+1}{n+1}\right) =$$

$$\sum_{q=0}^{n-k-1} \sum_{p=0}^q \binom{n+1}{q+1} \binom{q}{p} \binom{n+q-p}{q} (-1)^{q+p} \left(\frac{n-1}{n+1}\right)^p \left(\frac{n-q}{n+1} - \frac{k+1}{n+1}\right)^{n-p} (-1) =$$

$$\sum_{q=0}^{n-k-2} \sum_{p=0}^q \binom{n+1}{q+1} \binom{q}{p} \binom{n+q-p}{q} (-1)^{q+p} \left(\frac{n-1}{n+1}\right)^p \left(\frac{n-q}{n+1} - \frac{k+1}{n+1}\right)^{n-p} (-1) =$$

$$\sum_{p=0}^{n-k-1} \binom{n+1}{n-k+1} \binom{n-k}{p} \binom{n+n-k-p}{n-k} (-1)^{n-k-p} \left(\frac{k+1}{n+1}\right)^p \cdot 0^{n-p} (-1) = 0$$

Folglich ist in diesem Intervall:

$$F(x) = 1 + F_r(x) =$$

$$1 + \sum_{q=0}^{n-r-1} \sum_{p=0}^q \binom{n+1}{q+1} \binom{q}{p} \binom{n+q-p}{q} (-1)^{q+p+1} \left(\frac{n-1}{n+1}\right)^p \left(\frac{n-q}{n+1} - x\right)^{n-p}$$

4.4 Konsistenz des Tests

Ausgangspunkt der Arbeit war die Frage nach der "Konsistenz des Verfahrens". Abschließend kann nun dargestellt werden, daß der in Abschnitt 4.2 vorgeschlagene Test, eine geeignete Blockbildung vorausgesetzt, wirklich konsistent ist. Der Begriff der Konsistenz ist hier wieder in der schon in Paragraph 2 definierten Bedeutung zu verstehen. Wie schon angedeutet, sind zwei Voraussetzungen zu beachten:

- Ausgehend von einer Stichprobe $\vec{X}_1, \dots, \vec{X}_n$ unabhängiger und identisch verteilter p -dimensionaler Zufallsvektoren, sollen die StEB's so gebildet werden, daß der Vektor $(V_1, \dots, V_n) = (\mu_0(B_{k_1}), \dots, \mu_0(B_{k_n}))$ genau dann *Dirichlet* - verteilt ist, wenn μ_0 das tatsächlich zugrunde liegende Wahrscheinlichkeitsmaß war.
- Der Test, der auf den Zufallsvektor $(V_1, \dots, V_n)$ angewendet wird, muß konsistent sein.

Die Existenz einer Blockbildung, die den Erfordernissen von Punkt eins genügt und die entsprechenden Voraussetzungen, die dies gewährleisten, wurden in Paragraph 3 gezeigt. Der Einfachheit halber sei für die weiteren Betrachtungen eine Aufbereitung, wie in *Beispiel 3.5.7* vorgeführt wurde, zugrunde gelegt. Hierauf aufbauend betrachte man nur noch den Zufallsvektor $(V_1, \dots, V_n)$. B. Sherman¹⁾ konnte in seiner

1) Siehe Seite 64 unter 1).

Arbeit nachweisen, daß der von ihm entwickelte und in diesem Paragraphen verwendete Test konsistent ist. Seine Beweisführung gliedert sich in zwei Schritte:

Zunächst zeigt er, daß die Testgröße

$$S_n := S((V_1, \dots, V_{n+1})) = 1/2 \sum_{i=1}^{n+1} |V_i - 1/(n+1)|$$

asymptotisch normalverteilt ist, wobei $(V_1, \dots, V_n)$ Dirichlet - verteilt ist mit den Parametern

$\theta_0 = 1, \dots, \theta_n = 1$. D.h. die Zufallsgröße

$$\frac{S_n - E(S_n)}{\sigma(S_n)} \quad \text{ist asymptotisch standardnormalverteilt;}$$

dies geschieht, indem er nachweist, daß die Momente

der standardisierten Variablen $\frac{S_n - E(S_n)}{\sigma(S_n)}$ gegen die

entsprechenden Momente der $N(0;1)$ - Verteilung konvergieren.

Im zweiten Schritt folgert er aus dieser Aussage, daß es genügt, um die Konsistenz des Tests nachzuweisen, zu zeigen, daß der Erwartungswert $E(S_n)$ seiner Testgröße genau dann gegen

$e^{-1} = \lim_{n \rightarrow \infty} E(S_n)$ konvergiert, wenn der zugrunde liegende Zufallsvektor tatsächlich Dirichlet - verteilt war und die Varianz der Zufallsvariablen S_n bei allen Verteilungen des Hypothesenraumes mit wachsendem n gegen 0 konvergiert. Diese Aussagen werden von B. Sherman in seinem oben zitierten Artikel ausführlich bewiesen.

Zusammenfassend ist festzustellen, daß es eine Bildung von StEB's und einen darauf aufbauenden Test gibt, der es ermöglicht, mit wachsendem Stichprobenumfang eine falsche Hypothese immer besser zu erkennen.

§ 5

Abschließende Bemerkungen

In Paragraph 2 dieser Arbeit wurde gezeigt, daß unter gewissen Gesichtspunkten, vor allem dem der *eindeutigen Übertragung* des Hypothesenraumes Ω in eine Menge von Verteilungen der Gestalt

$$f(\vec{x}) = \begin{cases} c(\vec{x}) \cdot n! & , \text{ falls } 0 \leq x_i, i=1, \dots, n, \sum_{i=1}^n x_i \leq 1 \\ 0 & \text{sonst} \end{cases} \quad (*)$$

es genügt, nur die Projektionen auf die Koordinatenachsen als Schnittfunktionen ins Auge zu fassen. Es sind aber Prüfsituationen denkbar, in denen der Hypothesenraum, von der Sache bedingt, eingeschränkt ist, ohne daß das Problem schon "parametrisch" würde und somit nach einem prinzipiell anderen Test gesucht werden müßte; auch sind Fragestellungen möglich, die nur darauf abzielen, bestimmte Eigenschaften der Gestalt oder der Lage der Dichte der zugrunde liegenden Verteilung erkennen zu wollen. In diesen Fällen kann es sinnvoll sein, sich wieder anderer Schnittfunktionen als Projektionen zu bedienen, auch wenn hierdurch die Eindeutigkeit der Übertragung verloren ginge. Wie diese Funktionen auszusehen hätten, ist wohl theoretisch nicht entscheidbar, sondern müßte von Fall zu Fall empirisch ermittelt werden.

Wie in Satz 3.4.2 gezeigt wurde, haben die Dichten der transformierten Variablen die Gestalt (*), wenn eine Verteilung des Hypothesenraumes Ω zugrunde lag. Der in Paragraph 4 vorgeschlagene Test, wie auch die anderen Tests, die diesen Paragraphen einleitend in

Betracht gezogen wurden, eignen sich für Testsituationen, die alle stetigen Verteilungen, deren Dichten auf dem Simplex

$$\{\vec{x} \in \mathbb{R}_n : 0 < x_i, i=1, \dots, n; \sum_{i=1}^n x_i \leq 1\}$$

positiv sind, als Alternative zulassen. Interessant wäre ein Test, der auf den oben beschriebenen Fall spezialisiert ist, oder eine Untersuchung, wie sich der in Paragraph 4 vorgeschlagene Test bezüglich Alternativen dieser Art verhält. Diesem Problemkreis müßte man sich jedoch zunächst mittels numerischer Untersuchungen der Güteeigenschaften dieser Tests bezüglich verschiedener, in Betracht gezogener Alternativen nähern. Eine Abschätzung der Gütefunktion des in Paragraph 4 vorgeschlagenen Tests (in beide Richtungen) ist möglich, jedoch ohne numerische Hilfsmittel zu grob und wurde deshalb nicht in diese Arbeit aufgenommen. Eine weitere theoretische Vorbereitung, diese Fragen zu beantworten, wäre es, zu untersuchen, wie die Menge der in Frage kommenden Alternativen definiert werden könnte und welche Strukturen d.h. Metriken und Abstandsbegriffe auf dieser Menge definiert werden können, die dann eine möglichst gute Unterscheidbarkeit zur Hypothese (*Dirichlet* - Verteilung) zulassen.

Eine sich dieser Arbeit sofort anschließende Frage ist die nach einer Bereichsschätzung für die zugrunde liegende unbekannte Verteilung μ ; betrachtet man die Bildung der StEB's, so liegt es nahe, einen Konfidenzbereich für deren Verteilungsfunktion F_μ bzw. deren Dichte f_μ zu suchen. Ein Vorgehen in Dualität zu dem in Paragraph 4 vorgeschlagenen Test erscheint jedoch nicht sinnvoll: Geht man davon aus, daß der

Konfidenzbereich ein "Schlauch" sein sollte, der zu vorgegebenem Konfidenzniveau die "wahre" Verteilung d.h. ihre Verteilungsfunktion (oder Dichte) überdeckt; bedenkt man weiter, in Anlehnung an den ein-dimensionalen Fall, daß mittels der Blocks eine "empirische Verteilungsfunktion" mit dem Wertebereich $\{0, 1/(n+1), \dots, n/(n+1), 1\}$ konstruiert werden kann (n sei der Stichprobenumfang), so scheint das Aufsummieren

$$\sum_{i=1}^n |v_i - 1/(n+1)| + |n/(n+1) - \sum_{i=1}^n v_i|,$$

wodurch die Differenzen zu den Erwartungswerten der v_i zusammengefasst werden, nicht geeignet. Beispielsweise wäre es angebrachter, einen Test anzugeben, dessen Testgröße n -dimensional ist und auf dem Zufallsvektor $(V_1, \dots, V_n)$ beruht und dessen kritische Region im wesentlichen die Gestalt

$$K = \{\vec{v} \in \mathbb{R}_n : 0 \leq v_i, \sum_{i=1}^n v_i \leq 1, |v_i - 1/(n+1)| \geq \tau_\alpha, i=1, \dots, n\}$$

hat, wobei die τ_α durch die Gleichung $\mu_0(K) = \alpha$, wenn μ_0 die Verteilung der Hypothese bezeichnet, bestimmt sind. Eine ausführliche und systematische Behandlung dieser Fragestellung erscheint wünschenswert und interessant, würde aber über den Rahmen dieser Arbeit hinausführen.

Anhang zu § 4: 1)

Es werden nur die in diesem Paragraphen verwendeten Formeln angegeben

i) Exponential-Integral-Funktion

$$Ei(x) = \int_{-\infty}^x \frac{e^t}{t} dt \quad \text{für } x < 0$$

$$Ei(x) = -\lim_{\epsilon \rightarrow +0} \left| \int_{-\infty}^{-\epsilon} \frac{e^{-t}}{t} dt + \int_{\epsilon}^{\infty} \frac{e^{-t}}{t} dt \right| \quad \text{für } x > 0$$

ii) Sinus-Integral-Funktion

$$si(x) = - \int_x^{\infty} \frac{\sin t}{t} dt = - \frac{\pi}{2} + \int_0^x \frac{\sin t}{t} dt$$

Es gilt weiter

$$\lim_{x \rightarrow +\infty} (x^p si(x)) = 0 \quad \text{für } p < 1 \text{ und}$$

$$\lim_{x \rightarrow -\infty} si(x) = -\pi$$

iii) Cosinus-Integral-Funktion

$$ci(x) = - \int_x^{\infty} \frac{\cos t}{t} dt$$

iv) Es gilt

$$Ei(\pm ix) = ci(x) \pm isi(x)$$

1) Siehe Seite 102 unter 1), p. 925, p. 928 und p. 933.

Die *Leibniz'sche* Formel gibt die Ableitung n-ter Ordnung des Produktes zweier (differenzierbarer) Funktionen f und g an.

$$D^n(g \cdot f) = f \cdot D^n g + \sum_{i=1}^{n-1} \frac{n!}{(n-i)!i!} D^i f \cdot D^{n-i} g + D^n f \cdot g .$$

L I T E R A T U R - V E R Z E I C H N I S

ANDERSON, T. W.: Some Nonparametric Multivariate Procedures based on Statistically Equivalent Blocks. Multivariate Analysis. Proceedings of an International Symposium held in Dayton, Ohio, 1965. Edited by P. R. Krishnaiah, 5 - 27, Academic Press, New York and London (1966).

BILLINGSLEY, P.: Convergence of Probability Measures. J. Wiley & Sons, New York, London, Sydney, Toronto , (1968).

BILLINGSLEY, P., TOPSØE, F.: Uniformity and Weak Convergence. Zeitschrift für Wahrscheinlichkeitstheorie 7, 1 - 16, (1966).

FRASER, D. A. S.: Sequentially Determined Statistically Equivalent Blocks. Ann. math. Statistics 22, 372 - 381, (1951).

FRASER, D. A. S.: Nonparametric Tolerance Regions. Ann. math. Statistics 24, 44 - 55, (1953).

GÄNSSLER, P., STUTE, W.: Wahrscheinlichkeitstheorie. Springer-Verlag Berlin, Heidelberg, New York, (1977).

HALMOS, P. R.: Measure Theory. D. van Nostrand Company Inc., Princetown, New Jersey, Toronto, New York, London, (1950).

KEMPERMAN, J. H. B.: Generalized Tolerance Limits. Ann. math. Statistics 27, 180 - 186, (1956).

KIMBALL, B. F.: Some basic Theorems for developing Tests of Fit for the case of nonparametric Probability Distribution Function I. Ann. math. Statistics 18, 540 - 548, (1947).

MASSEY, F. J.: A note on the power of a nonparametric Test. Ann. math. Statistics 20, 440 - 442, (1950).

OLKIN, I., SOBEL, M.: Integral expressions for tail probabilities of the multinomial and negative multinomial distributions. Biometrika 52, 167 - 179, (1965).

RÉNYI, A.: Wahrscheinlichkeitsrechnung. VEB, Deutscher Verlag der Wissenschaften, Berlin (1966).

SHERMAN, B.: A random variable related to the Spacing of Sample values. Ann. math. Statistics 21, 339 - 361, (1950).

SCHUBERT, H.: Topologie. B. G. Teubner Stuttgart (1971).

TUKEY, J. W.: Nonparametric Estimation II, Statistically Equivalent Blocks and Tolerance Regions. The Continuous Case. Ann. math. Statistics 18, 529 - 539, (1947).

WALD, A.: An Extension of Wilks' Method for Setting Tolerance Limits. Ann. math. Statistics 14, 45 - 55, (1943).

WILKS, S. S.: Determination of Sample Sizes for Setting Tolerance Limits. Ann. math. Statistics 12, 91 - 96, (1941).

Als Formelsammlung wurde verwendet:

GRADSHTEYN, I. S., RYZHIK, I. M.: Table of Integrals, Series, and Products. Corrected and Enlarged Edition Prepared by A. Jeffrey. Acad. Press, New York, London, Toronto, Sydney, San Francisco (1980).

Lebenslauf

Name: Christina S c h n e i d e r

Geboren: 26. Dezember 1952 in Augsburg

Eltern: Richard Schneider, Oberamtsrat i. R.
Erna Schneider, geb. Pawellek

Familienstand: ledig

Schulen: 1959 - 1963 Volksschule in Augsburg
1963 - 1972 Anna Barbara v. Stetten'
sches Institut, Neusprach-
liches Gymnasium, Augsburg
23.06.1972 Reifeprüfung

Studium: WS 1972/73 - WS 1978/79
Mathematik an der Ludwig-Maximilians-
Universität München
18.12.1978 Diplom-Prüfung Mathematik

Berufliche Tätigkeit: Seit 01.02.1979 wissenschaftliche Ange-
stellte am Seminar für Spezialgebiete
der Statistik des Instituts für Stati-
stik und Wissenschaftstheorie der Uni-
versität München.

München, den 21. November 1984

